

OULUN YLIOPISTO  
Humanistinen tiedekunta  
Informaatiotutkimus

Jouni Pitkäaho

Paikallisuuden merkitys tiedon leviämisessä Twitterissä

Informaatiotutkimuksen

Kandidaatintutkielma

Oulu 2020

# SISÄLLYS

1 JOHDANTO .....	4
2 KESKEISET KÄSITTEET .....	5
2.1 Data .....	5
2.2 Sosiaalinen media.....	5
2.3 Sosiaalinen massadata .....	6
2.4 Aihetunniste .....	7
2.5 Tiedonlouhinta .....	7
2.5.1 Knowledge discovery in databases -prosessi .....	8
2.6 Paikkatieto.....	9
2.6.1 Paikkatietojärjestelmä .....	9
2.7 Twitter ja twiitit.....	9
2.8 Tiedon leviäminen .....	10
2.8 Viralisuus .....	10
3 TAUSTAA JA AIKAISEMPIA TUTKIMUKSIA .....	12
3.1 Tutkimusaiheen taustaa ja tutkimusilmiöön vaikuttavia tekijöitä .....	13
3.1.1 Informaatiokäyttäytyminen .....	14
3.1.2 Ympäristötekijät ja homofilia .....	15
3.1.3 Muut psykologiset ja sosiologiset tekijät .....	16
3.2 Aikaisemmat tutkimukset tiedon leviämiseen liittyen sekä tiedon kontekstin merkittävyys.....	18
4 MENETELMÄT, AINEISTONKERUU JA EETTISYYS .....	22
4.1 Menetelmät.....	22
4.2 Aineistonkeruu ja esittely.....	23
4.3 Tutkimuksen toteutus .....	25
4.4 Tutkielmassa käytetyt ohjelmistot .....	26
4.4.1 NCapture 1.0.290.0 .....	26
4.4.2 NVivo 12 .....	27
4.4.3 ArcMap 10.7.1.....	27
4.5 Tutkijan massadata -etiikka ja lainsäädäntö .....	27
4.5.1 Massadataan liittyvä tutkimusetiikka .....	28
4.5.2 Lainsäädäntö .....	28
4.5.3 Lisenssit ja luvat .....	29

5 TULOKSET .....	30
5.1 Ryhmän A tulokset .....	30
5.2 Ryhmän B tulokset .....	32
5.3 Ryhmän C tulokset.....	34
6 POHDINTA .....	36
7 LÄHTEET.....	39
8 LIITTEET .....	44
Liite 1. kartta .....	44
Liite 2. piirakkadiagrammi .....	45

## 1 JOHDANTO

Informaatiotutkimus on kirjava tieteenala, joka tutkii lähinnä ihmisen ja tiedon välistä vuorovaikutusta. Vaikka tietoa ja siihen liittyvää tiedon tallentamista, organisointia, luettelointia jne. on harjoitettu vuosisatojen ajan, niin alaa itsessään voidaan pitää nuorena (Mäkinen 2010, 24). Informaatiotutkimus on siirtynyt yhä etenevin määrin lähemmäksi sosiologiaa, psykologiaa ja tietojenkäsittelytiedettä, minkä vuoksi alaa voidaan pitää monitieteellisenä.

Digitalisaation myötä tiedon määrä on kasvanut räjähdysmäisesti. Internetin yleistyminen kotitalouksissa on johtanut tiedon määrän kasvamiseen ja internet, sekä sosiaalinen media, on antanut ihmisille mahdollisuuden ilmaista itseään, minkä vuoksi sosiaalisten medioiden tuottaman datan määrä on kasvanut merkittävästi (Bik & Goldstein 2013, 1). Datan määrän kasvaessa myös dataan liittyvät tutkimusmahdollisuudet lisääntyvät. Tässä tutkielmassa tarkastellaan hieman sosiaalisen massadatan maailmaa ja sen mahdollisuuksia.

Tutkielman pääaihe kohdistuu pääasiassa sosiaalisen median paikkaan liittyvän sisällön ja sisällön levittäjien paikkatietojen välisen suhteen tarkasteluun. Tutkielman tarkoitus on selvittää paikkaan sidottujen viestien leviämistä sosiaalisessa mediassa samalla kun otetaan huomioon viestien levittäjien geospaatialinen data, eli paikkatieto. Tutkielmassa halutaan selvittää miten paikalliset levittävät omaan paikkakuntaan liittyviä twiittejä verrattuna ulkopaikkakuntalaisiin.

Tutkielman toinen tarkoitus on toimia apuna tiedonlouhinnasta kiinnostuneille. Vaikka tutkielmassa ei käydä tarkemmin läpi tutkimusmenetelmää niin uskon, että keräämäni kattava kokonaisuus (ohjelmistot, tutkimukseen liittyvät tekijät jne.) saattavat auttaa tiedonlouhinnasta kiinnostuneita. Mielestäni massadata on yksi kiinnostavimmista yleistyvistä uusista asioista, minkä vuoksi näkisin massadatan aiheellisena ja ajankohtaisena asiana opettaa ja oppia.

## 2 KESKEISET KÄSITTEET

Tässä osiossa pyrin selittämään tutkielmassa esille nousseita käsitteitä. Vaikka joillekin käsitteille ja termeille ei ole vielä syntynyt vakiintunutta ja yleisesti hyväksyttyä määritelmää, niin olen pyrkinyt löytämään ja käyttämään tekstissä mielestäni kaikista tuoreimpia ja relevantimpia selityksiä kullekin käsitteelle. Toisin sanoen, käytän tutkielmassa puhekielessä ja opetuksessa käytettyjä käsitteitä ja niiden kuvailuja. Joidenkin aiheiden tutkiminen vaatii käsitteellistä analyysiä, jotta aiheen tulkinta ja havaintojen empiirinen keruu eivät jäisi merkityksiltään vähäisiksi (Tuomi 2007, 80). Aiheen paremman ymmärtämisen vuoksi käyn tässä osiossa läpi käsitteitä.

### 2.1 Data

Data on singulaarinen osa tietoa. Pelkästä yksilöllisestä datasta ei muodostu vielä tietoa, mutta suuresta määrästä dataa voidaan luoda tietoa ja ymmärrystä. Kotimaisten kielten keskus (2020) kuvailee dataa näin: ”vars. tietotekn. asian säännönmukainen esitys viestitettävässä t. käsittelykelpoisessa muodossa; tieto, tiedot.”

### 2.2 Sosiaalinen media

Sosiaalinen media on nimitys verkkoyhteisölle, jossa käyttäjät osallistuvat, vuorovaikuttavat ja jakavat itse tuottamia sisältöjä. Sosiaalinen media voi myös tarkoittaa yhteisöllisesti tuotettua ja jaettua mediasisältöä. (Heinonen 2009, 6.)

Sosiaalisen median voidaan ajatella olevan palvelu, jossa ihmiset harjoittavat verkossa tapahtuvaa sisällön vaihtoa. (Dewing 2012, 1). Vaikka suomen kielessä sosiaalisesta mediasta käytetään termiä ”yhteisöpalvelu” ja termin kuvauksena käytetään sanaa ”palvelua” (service), niin yhä useammin voi törmätä termiin ”alusta” (platform). Sosiaalinen media usein kattaa sisäänsä mobiili- sekä web-pohjaista teknologiaa luodakseen interaktiivisen alustan (Hermkens, Kietzmann, McCarthy & Silverstre 2001, 241). Termi ”alusta” pitää sisällään sosiaalisen median palveluiden muodostamaa

kokonaisuutta eri laitteistoilla, minkä vuoksi käytän tässä tutkielmassa käsitettä ”alusta” termin ”palvelu” sijaan.

Sosiaalisen median alustoina voidaan pitää mm. blogeja, wikeja, yhteisöpalveluita (Facebook, Twitter, Youtube jne.) (Heinonen 2009, 6; Dewing 2012, 1). Wikit ovat erilaisia servereitä, joilla säilytetään yleensä jonkin nettisivun tietokantaa, johon käyttäjät pääsevät käsiksi ja voivat osallistua sisällön luontiin (Wiki.org 2002). Tässä tutkielmassa keskitytään yhteen sosiaalisen median alustoista eli Twitteriin.

### 2.3 Sosiaalinen massadata

Massadata, tunnetaan myös termillä big data, tarkoittaa suurta määrää dataa, joka tyypillisesti seuraa kolmen V:n mallia: tiedon määrä (volume), muuttumisnopeus (velocity) ja monimuotoisuus (variety) (Laney 2001, 1). Massadata on usein sekavaa ja järjestäytymätöntä. Sosiaalinen massadata taas viittaa sosiaalisen median muodostamasta sisällöstä, jonka alustan käyttäjät ovat luoneet. Tässä tutkielmassa aineistona toimiva massadata on peräisin Twitterin käyttäjiltä. Massadataa on myös kuvailtu erilaisilla malleilla, mutta edellä mainittu malli on melko yleisesti käytössä.

Sosiaalisen median sisältö muodostuu monenlaisista eri mediatyypeistä kuten: tekstistä, videoista, kuvista, äänistä ja niin edes päin. Sisällön analysoimisessa koneellisesti vaikeutta tuo mm. se, että millä tarkoituksella tekstiä tuotetaan, eli onko tekstin kirjoittaja kritisoinut jotain vai onko viesti esimerkiksi sarkasmia. Kone ei ainakaan vielä osaa irrottaa viestin tarkoitukselta tai kontekstia tekstistä vaan tehtävä jää analysoinnista vastaavan henkilön vastuulle. Tässä tutkielmassa olemme kiinnostuneita enimmäkseen tekstimuotoisesta sisällöstä, eli twiittien aiheutunnisteista ja twiitin metatiedoista, lähinnä paikkatiedoista.

Sosiaalisen median sisältö voi sisältää oikean tiedon lisäksi disinformaatiota. Disinformaatio tarkoittaa tarkoituksellisesti vääristettyä tietoa tai tiedottamista (Valtiovarainministeriö 2008). Disinformaation jakaminen saattaa olla esim. poliittisesti tai kaupallisesti kannattavaa. Esimerkiksi öljy- ja tupakkayhtiöt ovat osallistuneet

disinformaation levittämiseen (Kolmes 2011, 34-35). Disinformaatiota voivat jakaa erilaiset tahot, kuten trollit, botit, valeutissivustot, valtiot, poliitikot, salaliittoteoreetikot ja mediat (Barberá ym. 2018, 22).

Disinformaation levittämisen mahdollisuus, ja sisällön tarkoituksien epäselvyys, on otettava huomioon data-analyysia tehdessä, sillä esim. tahallaan levitetty disinformaatio voi antaa väärän kuvan yleisestä mielipiteestä tai diskurssista. Toisaalta disinformaatio voi luoda uusia tutkimusaiheita. Tässä tutkielmassa ei sinällään tarvitse huomioida viestien sisältöä, mahdollista disinformaatiota tai viestien tarkoituksia, vaan tärkeimpänä elementtinä tutkimuksen kannalta pidetään paikkatietojen luotettavuutta.

## 2.4 Aihetunniste

Hashtag on puhekielessä tutumpi käsite kuin ”avainsana” tai ”aihetunniste”, mutta tässä tutkielmassa käytän käsitteistä jälkimmäistä. Aihetunniste on Twitterissä yleisesti käytetty tapa ilmaista tapahtumia, auttaa tuotteiden promotoimisessa tai helpottaa käyttäjien tapaa kategorioida twiittejään (Hoang, Kywe, Lim & Zhu 2012). Se merkitään ristikkomerkillä (#) ja jälkeisenä merkkijonona, joka sitoo tiedon tiettyyn teemaan tai aihepiiriin. Aihetunnistetta voidaan pitää metatietona, eli tietona tiedon sisällä. Tässä tutkielmassa sosiaalista massadataa analysoidaan osittain käyttäen paikkaan viittaavia aihetunnisteita.

## 2.5 Tiedonlouhinta

Tiedonlouhinta on menetelmä tai prosessi, jonka avulla datasta saadaan luotua tietoa. Tiedonlouhintaa voidaan myös kuvailla työkaluna, jonka avulla tietämys saadaan löydettyä datasta (Han, Kamber & Pei 2012, 1). Data voidaan kuvitella olevan epäjärjestyksessä olevina pikseleinä, joita järjestelemällä saadaan luotua kuva (ymmärrys) tilanteesta. Väärin järjesteltynä lopputuotos, eli kuva, voi näyttää hyvin

erilaiselta. Tiedonlouhintaa voidaan hyödyntää monipuolisesti eri tarkoituksessa, kuten luottokorttipetosten havaitsemisessa, sään ja rahoitusmarkkinoiden tilan ennustamisessa, markkinoinnissa tai tuotesuunnittelussa (Bramer 2007, 3).

#### 2.5.1 Knowledge discovery in databases -prosessi

Knowledge discovery in databases -prosessi (KDD-prosessi) on moniosainen tiedonlouhinnan prosessi, joka kuvaa tiedon ”louhimista” datasta. Prosessin vaiheita on kuvattu seuraavasti:

1. Datan siivoaminen, eli ei-halutun datan erottaminen pois halutusta datasta.
2. Datan integrointi, eli datalähteiden yhdistäminen.
3. Datan valitseminen, eli relevantin datan noutaminen tietokannoista (riippuen analyysin luonteesta).
4. Datan muuttaminen, eli datan muuttaminen muotoon, jolla sitä voidaan prosessoida.
5. Tiedonlouhinta eli prosessi, jossa käytetään tyypillisiä aiheeseen liittyviä menetelmiä.
6. Kuvioiden hahmottaminen, eli yhtälöiden tai mallien hahmottaminen myöhempää analyysia varten.
7. Tiedon esittäminen, eli tyypillinen presentaatio; tieto esitetään helposti ymmärrettävässä muodossa esim. kaavion avulla. (Han, Kamber & Pei 2012, 6.)

Knowledge discovery in databases -prosessi ei välttämättä aina etene edellä mainitun kaavan mukaan. KDD-prosessia voidaankin pitää enemmän suuntaan antavana mallina tiedonlouhinnasta kiinnostuneille. Kuten massadatan kolmen V:n malli, niin myös KDD-prosessilla on olemassa monia eri versioita. Tässä tutkielmassa käytetyn version voi sanoa koostuvan 5 askeleesta: 1. Datan etsiminen ja lataaminen, 2. Datan yhdistäminen, jos on tarvetta, 3. Datan muuttaminen haluttuun muotoon, 4. Datan käsittely, eli ”siivoaminen”, muokkaus ja analysointi niin kauan, kunnes 5. Datan palasista voidaan koostaa kuva, eli luoda ymmärrys.



## 2.6 Paikkatieto

Paikkatieto on yksinkertaisesti sanottuna tietoa, josta käy ilmi sijainti. Tämän tutkielman aineistossa paikkatieto on saatu twiittien geospatiaalisesta metadatasta. Twiittien paikkatieto ilmaantuu joko koordinaatteina (esim. +65.01236+025.46816/) ja/tai kirjoitetussa muodossa (esim. Oulu). Twitterissä oleva geospatiaalinen informaatio ei ole siis välttämättä niin tarkkaa, että siitä käy ilmi käyttäjien tarkka olinpaikka, mikä on yksityisyyden suojaamisen kannalta hyvä asia. Tässä tutkielmassa saatu paikkatieto saadaan aineiston Location -metatiedoista, ei koordinaateista.

### 2.6.1 Paikkatietojärjestelmä

Paikkatietojärjestelmä (GIS) on järjestelmä, jonka avulla paikkatietoa voidaan muun muassa analysoida, prosessoida, tallentaa ja esittää (Suomen Standardisoimisliitto 2020). Tässä tutkielmassa käytettävää paikkatietojärjestelmää käytetään ainoastaan tulosten visualisointiin.

## 2.7 Twitter ja twiitit

Twitter on suosittu sosiaalisen median alusta, jossa käyttäjät voivat jakaa maailmalle 280-merkin pituisia viestejä, eli twiittejä. Twiittiin voidaan lisätä kuvia, videoita tai gif -kuvia tekstin lisäksi. (Twitter 2020a.)

Twiitteihin voidaan myös lisätä linkkejä sekä aihetunnisteita, eli hashtagia. Twitterissä käyttäjät voivat seurata (follow) muita käyttäjiä, jolloin he näkevät seurattujen twiittejä omalla Twitterin etusivulla. Twitterin käyttäjät voivat myös ”uudelleenjakaa” muiden twiittauksia (retweet). Tässä tutkielmassa twiittauksilla viitataan twiittamiseen kuin myös retweettaamiseen.

Twitter on sosiaalinen verkosto. Sosiaalinen verkosto voidaan ajatella järjestelmänä, jossa järjestelmän käyttäjät ovat verkostoituneet keskenään ja ovat toisistaan

riippuvaisia pikemminkin kuin autonomisia yksilöitä (Faust & Wasserman 1994, 4). Twitterin käyttäjät ovat yhdistettynä toisiinsa monella eri tavoin ja heidän välillään vallitsevaa vuorovaikutusta voidaan tutkia.

## 2.8 Tiedon leviäminen

Tiedon leviäminen sosiaalisessa mediassa tarkoittaa erilaisten sisältöjen (viestien, kommenttien, kuvien, twiittien, videoiden, linkkien, mediasisältöjen jne.) leviämistä sosiaalisessa mediassa. Tiedon leviämisenellä on olemassa monta eri kuvaavaa termiä englanniksi, esim. spread of information, information spreading, information diffusion, going viral ja niin edespäin.

”Sosiaalinen verkosto” -termiä käytetään usein, kun puhutaan tiedon leviämisestä sosiaalisessa mediassa. Puhun tässä tekstissä yleisesti tiedon (tai informaation) leviämisestä (twiitit, sosiaalisen median sisältö, jne.), enkä erikseen valitse termiä tapauskohtaisesti; esim. en mainitse informaation diffuusiota (kun puhutaan sosiaalisesta verkostosta), tiedonkäyttöä (kun puhutaan informaatiokäyttäytymisestä) tai tiedon jakamista (kun tekstissä mainitaan tietojohtaminen). Tiedon leviäminen on mielestäni oiva yleistermi, joka yhdistää eri alojen termistöt saman termin alle mikä tekee tutkielman tekstistä helppolukuisempaa.

## 2.9 Viraalisuus

Akatemiaprofessori Hannu Salmi (2018) kuvaa viraalisuutta (viral) tapana, jolla jokin media leviää nopeasti internetissä ja saavuttaa laajan yleisön viruksen tavoin. Omasta mielestäni viraalisen statuksen saavuttanut sisältö ei aina välttämättä leviä nopeaa, mutta pidän henkilökohtaisesti kyseisestä termistä enemmän, kuin puhekielessä samaan asiaan viittaavista sanoista, kuten meemistä tai internetilmiöstä. Myös sanaa ”viraaliefektiä” käytetään jossain teksteissä, kuten esimerkiksi Mari Saaren opinnäytetyössä (2016).

Viralisuus -sana ei ole myöskään virallisesti suomen kieltä. Kyseinen termi on lähellä englanninkielistä vastinetta ”viral” niin siksi suosinkin sanaa ”viraalisuus” tässä tekstissä. Mielestäni termi itsessään on erinomainen, sillä se on osuva, joustava sekä se on läheinen englanninkielisen vastineen kanssa. Kaikissa yksinkertaisuudessaan määrittäisin itse viraalisuuden tilaksi, jossa teksti, uutinen, kuva, video, ääni (tai mikä tahansa mediamuoto) leviää laajalle.

### 3 TAUSTAA JA AIKAISEMPIA TUTKIMUKSIA

Tässä tutkielmassa pyritään löytämään vastaus kahteen kysymykseen:

1. Miten twiitit, jotka ovat sidottu kontekstiltaan tiettyyn paikkaan, leviävät Twitterissä viestiä levittävien käyttäjien paikkatietoon nähden?
2. Miten suuri vaikutus käyttäjien fyysisellä sijainnilla on siihen, miten aktiivisesti he ovat levittäneet paikkaan sidottuja twiitteja?

Tutkimuksessa yritetään selvittää leviävätkö paikkaan sidotut twiitit lähinnä kyseisessä paikassa asuvien Twitter-käyttäjien keskuudessa valtakunnallisella tasolla. Tutkimuksessa yhdistyy siis virtuaalinen sekä fyysinen maailma.

Käyn tässä osiossa läpi tutkimusaiheen taustaa ja varteenotettavia tekijöitä monen erin alan näkökulmasta. Ihmisen käyttäytymistä pohditaan yksilön, yhteisön ja verkoston tasolla. Tutkielmassa käydään myös läpi lyhyesti ympäristötekijöitä ja homofilian vaikutusta ihmisen käyttäytymiseen sosiaalisessa mediassa.

Pureudun myös aikaisempiin tutkimuksiin sekä pohdin tiedon kontekstin tärkeyttä tiedon leviämisen kannalta. Olen valinnut tähän kappaleeseen mielestäni tutkielman kannalta tärkeimmät tekijät. Aiheen kompleksisen luonteensa vuoksi, myös tutkimuksen ilmiöön vaikuttavat tekijät ovat osittain komplekseja ja monimuotoisia. Tässä osiossa on esitetty joitain tiedon leviämiseen vaikuttavia tekijöitä kirjallisuuskatsauksen avulla. Kirjallisuuskatsauksen tarkoitus on tuoda esiin näkökulmia aiheeseen liittyen aikaisempien tutkimuksien avulla (Tuomi 2007, 82). Tutkielmani aiheesta ei löydy juurikaan suoraan tutkimusartikkeleita, joten jouduin hakemaan lähteitä aiheen ”ympäriältä”. Tekijät ja lähteet on siis valittu johdonmukaisen päättelyn perusteella; tieteellisiä artikkeleita on verrattu toisiinsa ja parhaimmat on valittu. Toisin sanoen, tekijät ja artikkelit on valittu harkinnanvaraisesti ja humanistisiin tieteenaloihin nojautuen.

Kaikkien tutkielman artikkeleiden, tekijöiden ja käsitteiden on tarkoitus palvella lukijaa antamalla hänelle mahdollisemman laaja ymmärrys monimuotoiseen aiheeseen liittyen.

Mahdollisemman laajan kuvan ansiosta tutkittavaa ilmiötä, sekä tutkimustuloksia, on mahdollista ymmärtää paremmin.

### 3.1 Tutkimusaiheen taustaa ja tutkimusilmiöön vaikuttavia tekijöitä

Twitter, kuten muutkin sosiaaliset mediat, ovat kasvaneet internetin yleistymisen sekä teknologisen kehityksen myötä. Internetistä on tullut työkalu, jonka avulla ihmiset voivat mm. pitää yllä tai luoda uusia sosiaalisia kontakteja, tuoda omia mielipiteitä esiin tai kuunnella toisten mielipiteitä. Twitterin alustalla vieraili 2019 neljännellä kvartaalilla keskimäärin 152 miljoonaa aktiivista päivittäistä käyttäjää, joille voitiin näyttää mainoksia, eli heillä voitiin ansaita rahaa (Twitter 2020b, 2). Todellinen kävijämäärä on todennäköisesti todellisuudessa enemmän. Twitteriä hyödyntävät yritykset, poliittiset tahot, yksilöt, organisaatiot sekä monet muut osallistujat. Olipa kyse sitten markkinoimisesta, tiedottamisesta tai viestinnästä, niin Twitter on luonut itselleen jalansijan sosiaalisen medioiden maailmassa. Twitter ja sen käyttäjät luovat monenlaisia eri ilmiöitä, joita monet eri tieteenalat voivat tutkia.

Tutkielman tutkimusosuudessa ei keskitytä siihen, miksi tai miten ihmiset jakavat tietoa sosiaalisessa mediassa, vaan miten paikkaan sidottu tieto leviää. Toisin sanoen, ihmisen informaatiokäyttäytymiseen sekä muihin psyko-sosiologisiin tekijöihin ei ole tarpeellista uppoutua liiakseen. Käyn tosin tässä osiossa lyhyesti läpi mahdollisia tekijöitä, jotka voivat vaikuttaa tuloksiin, sillä näiden tekijöiden huomioon ottaminen voi auttaa meitä ymmärtämään tutkielmassa tutkittavia ilmiötä paremmin.

Vuonna 2017 julkistetun tutkimuksen mukaan 7-15% Twitterin käyttäjistä on botteja (Davis, Ferrara, Flammini, Menczer & Varol 2017). Voidaan varovaisesti olettaa, että suurin osa kyseisen alustan käyttäjistä ovat ihmisiä tutkielman teon aikana 2020. Tietoa siis levittää ja vastaanottaa pääasiassa ihminen. Siispä tässä osiossa keskitytään ihmiseen ja ihmisen käyttäytymiseen. Todella moni eri tekijä vaikuttavaa siihen, miten ja miksi tietoa jaetaan, sekä miten ihminen käyttäytyy sosiaalisessa mediassa. Olen valinnut mielestäni oleellisimpia tekijöitä eri tieteenaloilta.

### 3.1.1 Informaatiokäyttäytyminen

Informaatiokäyttäytyminen on yksi informaatiotutkimuksen osa-alueista. Informaatiokäyttäytymisessä tutkitaan lähinnä ihmisen tapaan käyttää informaatiota, siitä nimi ”informaatiokäyttäytyminen”. Informaatiokäyttäytymisestä tutut teoriat ja tutkimukset toimivat mielestäni hyvänä alustana, tai pohjana, tutkielman ilmiön tarkastelun kannalta.

Informaatiokäyttäytyminen käsittelee mm. ihmisen tiedonhakua, tiedon vastaanottamista sekä tiedonkäyttöä (Wilson 2000, 49-50). Tiedonkäyttöön on artikkelin mukaan yhdistynyt sekä fyysisiä, että henkisiä toimia. Ihmisten kognitiiviset, affektiiviset, sosiaaliset, kulttuuriset, kielelliset sekä monet muut eri tekijät voivat vaikuttaa henkilön informaatiokäyttäytymiseen (Bruce, Fidel & Pettigrew 2001, 68). Jo negatiiviset tunteet voivat muuttaa ihmisen tietokäyttäytymistä, aiheuttaen tiedonvälttelyä (Sairanen & Savolainen 2008, 6). Ihmisillä on eri arvot, maailmankuvat ja ennakkoluulot, mitkä voivat myös vaikuttaa tiedon jakamiseen; poliittisen spektrin oikealla tai vasemmalla puolella arvomaailmaan uskova yksilö luultavasti jakaa mielellään twiittejä, jotka jollain lailla vahvistavat hänen maailmankuvaansa, eivätkä ole sitä vastaan. Poliittisia viestejä levitetään Twitterissä nopeammin ja huomattavasti enemmän, jos ne herättävät käyttäjässä positiivisen tunteen (Dang-Xuan & Stieglitz 2013, 239-241). Minusta on selvää, että huvittavat ja hauskat asiat vaikuttavat sisältöjen leviämiseen, varsinkin kun puhutaan viraallisista sisällöistä. Positiivinen tunne on mielestäni yksi tiedon leviämistä motivoiva tekijä. Omakohtaisien kokemusten perusteella kontekstiltaan hauskat viestit yleensä leviävät parhaiten.

Chatmanin (1991; 1996) ”pieni maailma” -teoria (small world) sekä Haasion (2015, 2) kirjoittama artikkeli osoittavat, että ihmisillä on tapana jakaa ja vastaanottaa tietoa riippuen heidän ympäristöstään. Yksilön sekä yhteisön asettamat normit ja arvot säätelevät informaatiokäyttäytymistämme (Haasio 2015, 2). Se, mitä tietoa jaetaan, voi riippua hyvin paljon siitä onko se ”hyväksyttyä” tietoa, eli vastaako se yhteisön tai yksilön normeja ja arvoja.

Sosiaalisen median käyttäjät voivat luoda itsensä ympärille siis oman maailmansa, jossa jaetaan lähinnä hänen arvoihinsa ja normeihin nojautuvaa sisältöä. Se, mitä tietoa jaetaan ja vastaanotetaan sosiaalisessa mediassa, voi siis riippua hyvin paljon siitä millaisen ympäristön käyttäjä on luonut itselleen. Oman ympäristön voi luoda sosiaalisessa mediassa mm. seuraamalla tiettyjä ihmisiä, jonka jälkeen heidän julkaisemat viestit näkyvät ensisijaisesti heidän seuraajilleen tai liittymällä erilaisiin ryhmiin. Käyttäjät voivat myös blokata muita käyttäjiä, joiden viestejä ei halua nähdä. Ympäristöistä kerrotaan enemmän omassa kappaleessa 4.1.2 (Ympäristötekijät ja homofilia).

Ihmisillä on paljon eri tuntemuksia ja tapoja reagoida asioihin. On olemassa paljon kognitiivisia, kuin myös affektiivisia tekijöitä esim. tunteita, tuntemuksia, mielialoja yms., jotka säätelevät informaatiokäyttäytymistä tiedonhaun prosessissa (Kuhlthau, 2020). Se, missä määrin mikäkin affektiivinen tekijä vaikuttaa eniten tiedon jakamiseen sosiaalisessa mediassa, on hankala määritellä. Tämän (sekä aiheen monimutkaisuuden) vuoksi en uppoudu sen syvemmälle affektiivisiin ja kognitiivisiin tekijöihin tässä tutkielmassa.

Sosiaalisen median käyttäjien kyky ymmärtää viestien sisältö voi myös osoittautua tiedon jakamisen kannalta tärkeäksi tekijäksi. Informaatio-, media-, sekä muut lukutaidot liittyvät myös jossain määrin ihmisen informaatiokäyttäytymiseen. Uskoisin, että tietoa on hankala käyttää (esim. jakaa) ja sen arvoa on hankala ymmärtää, jos yksilö ei ymmärrä tai osaa arvioida kyseisen tiedon sisältöä.

### 3.1.2 Ympäristötekijät ja homofilia

Tutkimustuloksia ymmärtäessä olisi myös hyvä ottaa huomioon muutkin ympäristötekijät, kuten esimerkiksi alusta tai nettisivu ja niihin liittyvät seikat; emme tiedä Twitterin käyttämiä koodeja ja algoritmeja, joten emme voi olla varma minkä perusteella tiettyjä twiittejä esitetään kullekin käyttäjälle tai miksi jotkin twiitit ovat toistaan näkyvämpiä. Myös käyttäjäkunta saattaa vaikuttaa alustalla käytyyn poliittiseen

diskussioon ja siihen mitä on hyväksyttyä jakaa ja mitä ei. Jos seuraa vain tietyn poliittisen arvomaailman omaavia käyttäjiä, niin luultavasti myös saa nähdä normaalia enemmän kyseiseen poliittiseen arvomaailmaan liittyviä twiittejä. Homofilia on sosiaalipsykologian termi, joka tarkoittaa yksilöiden mieltymystä samankaltaisuutta kohtaan (Tieteen Termipankki 2015). Homofilialla on tärkeä rooli tiedon leviämisen suhteen sosiaalisessa mediassa, sillä ihmiset keskustelevalt verkossa enemmän muiden henkilöiden kanssa, jotka ovat heidän kaltaisiaan (Zinoviev 2011, 8).

### 3.1.3 Muut psykologiset ja sosiologiset tekijät

Ihmisen käyttäytymiseen sosiaalisessa mediassa vaikuttavat myös psykologiset - että sosiologiset tekijät. Tämän vuoksi, tässä kappaleessa käydään syvemmin läpi tekijöitä, jotka voivat myös vaikuttaa tutkittavaan ilmiöön.

Muita sosiaalisen median tiedon jakamiseen liittyviä käyttäytymistä ohjaavia tekijöitä voisivat olla esimerkiksi erilaiset kognitiiviset vinoutumat (cognitive biases). Kognitiivisilla vinoutumilla tarkoitetaan tapaa havaita maailmaa tietynlaisten linssien läpi mikä voi johtaa mm. sosiaalisen tiedon väärinkäyttöön (Funder & Krueger 2004, 1). Erilaisia kognitiivisia vinoutumia on iso määrä, mutta mielestäni vahvistusharha (confirmation bias) on niistä tärkein, tai yksi tärkeimmistä, kun keskitytään tiedon leviämiseen. Vahvistusharha kuvaa tilannetta, jossa henkilö suosii tietoa, joka vastaa hänen ennakkoluulojaan samalla vahvistaen hänen näkemyksiään (Anttila, Jokinen & Ruponen 2018, 42; Ahola, Bäckström & Mäkinen 2014, 118). Henkilö siis voi valita, mitä tietoa hän jakaa riippuen siitä millaisia ennakkoluuloja hänellä on kyseistä tietoa kohtaan.

Sosiaalipsykologian näkökulmasta ihmisten verkkovuorovaikutukseen liittyy monia tekijöitä aina anonyymiuudesta luovuuteen, kuin myös yksilön tasosta joukkojen tasoon (Matikainen 2008, 25-30). Matikaisen (2008, 32) mukaan ryhmään samaistuminen on tärkeä prosessi, sillä silloin kun yksilö tuntee ryhmää kohden yhteenkuuluvuutta niin hän myös saattaa viestiä ystävällisemmin. Ihminen voi käyttää aggressiivisempaa kieltä, kun



hän tuntee olevansa ryhmässä (esim. jollain keskustelufoorumilla) johon, ja jonka käyttäjiin, hän ei tunne minkäänlaista yhteyttä. Twitter ympäristönä saattaa siis tuoda aggressiivisen puolen esiin joillakin käyttäjillä, jos he eivät tunne kuuluvansa sinne.

Internetin tarjoama ”suoja”, eli anonyymius, muokkaa myös ihmisen käyttäytymistä internetissä. Anonyymisuojan tiedetään lisäävän aggressiivisuutta internetympäristössä (Zimmerman & Ybarra 2016, 10). Anonyymiuden tiedetään myös lisäävän verkkokiusaamista, sillä kiusaaja harvoin kohtaa negatiivisia seuraamuksia (Barlett 2015, 75). Tiedon jakamiseen liittyvä käyttäytyminen sosiaalisessa mediassa esim. twiittaus tai retwiittaus voivat poiketa hyvinkin siitä, miten yksilö käyttäytyisi normaalisti verkon ulkopuolella, jos hän toimisi kasvotusten tai ei-anonyymisti. Sosiaalisen median käyttäjä uskaltaa jakaa kiisteltyä tai keskustelua herättävää sisältöä internetissä vapaammin, sillä hän saattaa tuntea, ettei hänen teoillaan ole negatiivisia seuraamuksia. Anonyymius, ja pelkästään katsekontaktin puuttuminen, voivat lisätä negatiivista ilmapiiriä tai uhkailun määrää verkkoviestinnässä (Barak & Lapidot-Lefler 2011, 439).

Vahvistusharha sekä anonyymius on mainittu tässä tekstissä sen vuoksi, koska yksi tutkielmassa käydyistä aihetunnisteista (#TurvallinenOulu) loi paljon kiivasta poliittista keskustelua Twitterissä. Kiivaat poliittiset nettikeskustelut ovat osittain aggressiivisia ja omien kokemuksieni, havaintojeni, sekä aikaisempien tutkimuksien pohjalta voin vetää varovaisen johtopäätöksen, että nämä kaksi em. tekijää voivat auttaa kyseisen aihetunnisteen liittyvän aineiston tutkimisessa.

Tästä kappaleesta on jätetty pois joitain mahdollisia tekijöitä, kuten tietojohdamisesta tutut tiedon jakamiseen liittyvät teoriat ja mallit, sillä ne painottuivat mielestäni lähinnä yritysmailmaan ja organisaation sisälle. Jätin myös pois erilaisia filosofiasta tuttuja tekijöitä, jotka liittyivät ihmisten tapaan kommunikoida, sillä ne olivat minusta liian teoreettisia. Olen myös jättänyt pois kaikki ulkoiset tekijät (myös: ulkoiset vaikuttajat, external influence) esim. perinteisen median. Perinteinen media on yleensä se, joka uutisoi ilmiöistä, ei luo niitä. Tutkimuksessa käytetyt aineistojen avainsanat eivät ole saaneet itselleen kunnollista mediahuomiota tutkielman teon aikana, joten en näe

tarpeelliseksi tuoda esiin artikkeleita median vaikutuksesta tiedon leviämiseen sosiaalisessa mediassa.

Tiedon leviämisestä löytyy monelta eri tieteenalta erilaisia näkemyksiä, jotka muistuttavat ja samalla poikkeavat toisistaan hyvinkin paljon. On mahdotonta mahduttaa tähän tutkielmaan kaikkien eri alojen omia näkemyksiä aiheeseen liittyen. Siksi valitsin omasta mielestäni relevanttimmat ja uusimmat näkemykset, joista näin olevan parhaiten hyötyä tutkielman aiheen, tuloksien sekä ilmiön ymmärtämisen kannalta.

### 3.2 Aikaisemmat tutkimukset tiedon leviämiseen liittyen sekä tiedon kontekstin merkittävyys

Tiedon leviämistä sosiaalisessa mediassa on tutkittu jonkin verran. Yleisesti ottaen tiedon leviämiseen vaikuttaa hyvin paljon esimerkiksi tiedon konteksti; juoruista koostuvaa materiaalia levitettiin sosiaalisessa mediassa suuremmalla määrällä ja paljon tarkemmin mitä (ei-sosiaalista) yksityistä ja fyysistä tietoa (Dunbar, Mesoudi & Whiten, 412-413). Tiedon kiinnostavuus siis voi olla yksi merkittävä tekijä, kun tarkastellaan tiedon levittämistä sosiaalisessa mediassa.

Paikallisuus on mielestäni yksi tärkeimpiä konteksteja tiedon jakamisen osalla. On sanomattakin selvää, että toisella puolen maapalloa tapahtuneet asiat eivät välttämättä puhuta suomalaisia ja toisin päin. Paikalliset uutiset koskettavat meitä enemmän, sillä silloin tiedämme, että me tai lähipiirimme ovat mahdollisesti altistuneena uutisoitavalle ilmiölle. Paikallisuuden herättämät tunteet ovat luultavasti yksi pääsyistä siihen, miten ja mitä tietoa jaetaan ja kuinka paljon.

Paikallisuuden lisäksi uskon, että tutkielmani ilmiöön vaikuttaa myös ihmisten informaatiokäyttäytymistä ohjaava itsekeskeisempi tekijä kuten esimerkiksi niin kutsuttu self-reference effect. Self-reference effect on ilmiö, jossa ihminen käsittelee informaatiota liittyen siihen, miten paljon kyseinen informaatio vaikuttaa heidän omaan

elämäänsä; ihminen esimerkiksi muistaa paremmin tietoa, joka liittyy häneen itseensä (Kihlstrom & Klein 1986, 34; Johnson & Symons 1997, 377).

Tutkijat ovat seuranneet informaation leviämistä monilla eri alustoilla. Twitterille on kehitelty erilaisia malleja, joiden avulla twiittien leviämistä voidaan mallintaa (Kawamoto 2013, 34). Informaatio leviää eri tavalla eri alustoilla (Hatano & Kawamoto 2014). Informaation leviämiseen eri alustoilla vaikuttaa monet seikat, kuten aiemmin on mainittu (mm. algoritmit ja käyttäjäkunta). Hatanon ja Kawamoton (2014) mukaan sosiaalisen median sisällöllä on myös olemassa ”kallistuspiste” (tipping point), joka määrittelee sen, milloin sosiaalisessa mediassa jaetusta viestistä tulee viraali. Heidän tekemän tutkimuksen mukaan kallistuspiste on eri riippuen alustasta. Tämä tarkoittaa sitä, että jollain alustoilla sisältö voi muuttua nopeammin viraaliksi.

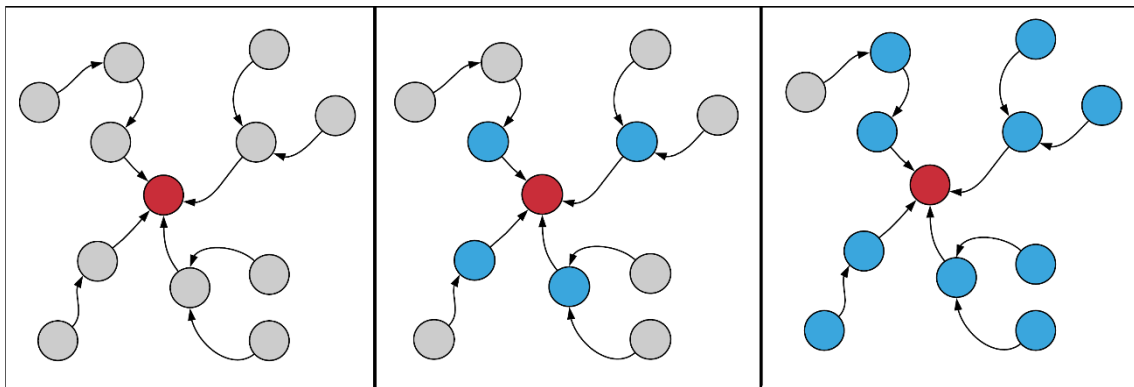
Sosiaalisen median (sosiaalisen) verkoston vuoksi käyttäjien toiminnot (sosiaalisessa mediassa) vaikuttavat muihin käyttäjiin – tästä käytetään termiä ”sosiaalinen vaikuttaminen” (Favre, Guille, Hacid & Zighed 2013; Tieteen termipankki 2014). Sosiaalinen vaikuttaminen tarkoittaa käytännössä sitä, että yksilö toimii ympäristön (esim. muiden käyttäjien) mukaisesti, kun hän jakaa sosiaalisen median sisältöä eteenpäin esim. retweettaamalla. Favren, Guillen, Hacidin ja Zighedin (2013) artikkelin mukaan vaikutusvaltaiset levittäjät (influential spreaders) on otettava huomioon, kun tarkastellaan tiedon leviämistä sosiaalisessa mediassa. Artikkelissa myös huomioidaan, että 71% Twitterin informaation liikehdinnän määrästä tapahtuu Twitterin sisäisten vaikutuksien voimasta, kun taas 29% ulkoisten tapahtumien ja vaikutuksien voimasta.

Sosiaaliset linkit (social links) ovat hallitsevin tekijä tiedon leviämisen kannalta sosiaalisessa medioissa (Fischer & Taxidou 2014). Tieto leviää sosiaalisten linkkien kautta henkilöltä toiselle. Kuviossa 1 käy ilmi, miten vaikutusvaltaisten levittäjät (sosiaalisen median käyttäjiä, joilla on paljon seuraajia, kavereita tmv.) vaikuttavat tiedon leviämiseen.

Tiedon leviämistä sosiaalisessa mediassa on pyritty kuvaamaan eri tavoin. Yksinkertainen tapa on esitetty Kuviossa 1, jossa punainen pallo kuvastaa henkilöä, joka on tuottanut sisältöä. Harmaat pallot kuvastavat henkilöitä, jotka ovat ”linkitettyinä”

muihin henkilöihin, mutta eivät ole vielä kontaktissa sisältöön. Siniset pallot kuvastavat sitä, missä vaiheessa linkin seuraava jäsen on jakanut sisällön eteenpäin (esim. twiittaamalla tai retweettaamalla). Tiedon leviäminen tapahtuu mallissa askel askeleelta ja mallin esimerkeistä 100% päättää jakaa sisältöä eteenpäin. Nuolet kuvastavat henkilöiden välisiä suhteita; nuolen pää osoittaa sitä henkilöä, jota hän seuraa eli henkilö näkee seurattujen jakamaa (esim. retweetattua) sisältöä.

Kuvio 1. Yksinkertaistettu malli tiedon leviämisestä sosiaalisessa mediassa.



Maantieteellisellä sijainnilla on merkitystä tiedon leviämisen kannalta sosiaalisissa verkostoissa. Sosiaalisen median käyttäjillä on vähemmän pitkän matkan kontakteja (long range connections) (Currie ym. 2012, 23). Tämä on ymmärrettävää, koska on normaalia luoda ystävyysuhteita paikallisesti esim. koulussa tai työpaikoilla. Myöhemmin nämä ystävät saatetaan lisätä kontakteihin sosiaalisessa mediassa. Toisaalta juuri Twitterin käyttäjillä oli yhteydet luotu muihin käyttäjiin etäisyyksiltään monipuolisesti verrattuna muiden alustojen käyttäjiin (Latora, Mascolo, Musolesi, & Scellato 2010). Artikkelin spekuloi, että Twitterin käyttäjiä luultavasti kiinnostaa ystävystyä käyttäjien kanssa, jotka jakavat mielenkiintoista sisältöä pikemminkin kuin pelkästään lähietäisyyksien vuoksi.

Palenin ja Starbirdin (2012) tutkimuksessa havaittiin, että twiittejä jaetaan eri tavalla eri paikassa. Artikkelissa tarkasteltiin Egyptin vallankumouksen aikaista Twitterissä tapahtuvaa informaatioliikennettä. Twiittejä tutkimukseen valittiin aihetunnisteilla #egypt ja #jan25 ja sanalla egypt. Tutkimuksen otoskappaleen avulla huomattiin, että 29,8% otoksen käyttäjien olinpaikasta oli pystytty yhdistämään Kairoon, 6% muuhun

egyptiin, 17,3% arabimaihin ja 45,6% arabimaiden ulkopuolelle. Tutkimuksessa huomautettiin, etteivät he pystyneet paikantamaan yli 800 000 Twiittaaajaa, mutta he epäilivät, ettei yksi kolmasosa heistä olisi voinut olla paikan päällä Kairossa. Heidän tutkimuksensa voi antaa tietoa siitä, miten paikkaan viittaavien avainsanojen merkitys näkyy tiedon leviämisessä levittäjien paikkatietoihin nähden. Toisaalta on hyvä huomioida muutamia seikkoja: Ensinnäkin, vain n. 25,6 prosenttia Egyptin kansalaisilla oli pääsy internettiin vuonna 2011 (The World Bank 2020). Internet yleistyy kehittyneillä alueilla ensiksi esim. isoissa kaupungeissa, joten alhaisen käyttäjäprosentin perusteella voidaan olettaa, ettei internet ei ollut vielä kovin yleinen Kairon ulkopuolella. Tämän voisi selittää Kairon korkean prosenttiosuuden (29,8%). Toiseksi, aihetunnisteet #egypt ja #jan25 sekä sana ”egypt” ovat englannin kielellä, mikä saattaa edesauttaa kyseisten aihetunnisteellisten twiittien leviämistä maissa, joissa Twitterin käyttäjät osaavat englantia. Otoksesta 54/248 (n. 22%) käyttäjistä olivat kotoisin Pohjois-Amerikasta. Lukuihin on kuitenkin syytä suhtautua kriittisesti, sillä otoksesta vain kuusi prosenttia oli Kairon ulkopuolella asuvia egyptiläisiä. Mielestäni prosenttiosuus kuulostaa liian vähältä.

Näiden seikkojen perusteella voidaan olettaa, ettei Palinin ja Starbirdin (2012) tutkimuksen perusteella voida vielä täysin päätellä, miten paikkaan liittyvien aihetunnisteiset viestit leviävät sosiaalisessa mediassa. Heidän tutkimuksensa luettua huomasin, että valtioiden internetin käyttäjämäärät, kansalaisten kieliosaaminen sekä eri kielisten aihetunnisteiden käyttö ovat varteenotettavia seikkoja ja ne olisi hyvä ottaa huomioon oman kandidaatintutkielman teossa. Suomessa ei ole syvää teknologista kuilua, aihetunnisteet ovat suomeksi ja tutkielmassa keskitytään Suomen sisäiseen verkkoviestintään.

Tiedon tai sosiaalisen median sisällön kontekstisuus voi vaikuttaa tiedon leviämiseen hyvinkin monen tavoin. Ilmiönä tiedon leviäminen on sangen monimutkainen, mutta monimutkaisuus taas tekee aiheesta monipuolisen tutkittavan. Aiheesta syntyy varmasti paljon tutkimusmahdollisuuksia tulevaisuudessa.

## 4 MENETELMÄT, AINEISTONKERUU JA EETTISYYS

Tässä luvussa käsittelen tutkimukseni menetelmiä, kuvaan aineistonkeruuta ja tutkimuksen toteuttamista. Lopuksi käydään läpi tutkimuksessa käytettyjä ohjelmia sekä pohdin tutkimuksen tekemiseen liittyviä eettisiä kysymyksiä.

### 4.1 Menetelmät

Tarkastelen tutkielman ilmiötä empiirisesti, tutkien kahden tekijän suhteen riippuvuutta toisiinsa; paikkaan viittaavan tiedon ja kyseisen tiedon levittäjän paikkatiedon välistä suhdetta. Tutkielman lähestymistapa on eksploratiivinen, sillä tutkimuksen ilmiötä ei ole aiemmin juurikaan tutkittu. Tutkielma ei siis perustu mihinkään teoriaan. Näkökulman on tarkoitus muodostua tutkielman teon aikana.

Eksploratiivisella tutkimuksella tarkoitetaan tuntemattoman tai vähän tutkitun tutkimuskohteen tutkimista. Empiirisellä aineistolla luodaan hypoteeseja ja teoreettisia ideoita sen sijaan, että niitä koeteltaisiin. Ilmiötä voidaan kuvata teoreettisesti empiiristen havaintojen avulla. Eksploratiivinen tutkimus ei aina anna vastausta tutkimusongelmaan, mutta se mm. antaa vihjeitä siitä, mistä vastauksen voi löytää. (Uusitalo 1991, 62.)

Aineiston aihetunnisteet valittiin harkinnanvaraisesti. Twitterin algoritmit määrittelivät mitä twiittejä annettiin ladattavaksi. Twitterin algoritmit eivät ole julkista tietoa, mutta voidaan olettaa, että ne noudattavat jotain logiikkaa. Tutkielmassa voidaan siis puhua näytteistä otoksien sijaan. Datan keruuseen vaikuttavista tekijöistä kerrotaan tarkemmin kappaleessa 5.2 (aineiston esittely ja -keruu). Aihetunnisteiden valintoihin vaikutti lähinnä oma mielenkiinto ja mahdollisuus niiden keskinäiseen vertailuun.

Aihetunnisteet valittiin harkituin perustein: #Oulu ja #TurvallinenOulu valittiin, koska molemmat liittyvät Ouluun. Molempia näytteitä voidaan verrata keskenään, sillä ne ovat paikkatiedoltaan samoja, mutta kontekstiltään eriäviä. Oletan, että kontekstin eriävyys näkyisi jollain tavalla myös tuloksissa. Aihetunniste #Helsinki valittiin sen vuoksi, koska Helsinki on Suomen suurin paikkakunta ja mielestäni Helsinki on myös maailmalla

tunnetuin suomalainen kaupunki. Halusin nähdä miten maailmanlaajuinen tunnettavuus vaikuttaa tuloksiin. #Porvoo valittiin, sillä Porvoo on etäisyydeltään lähellä Helsinkiä ja halusin nähdä, miten Porvooseen liittyvä aihetunniste leviää väkiluvultaan isomman kaupungin vieressä. #UusimaaLockDown -aihetunniste valittiin siksi, että se oli tutkielman teon aikana ajankohtainen ja siitä saatiin iso määrä twiittejä ja retwiittejä. #kissmyturku -aihetunniste valittiin, koska kyseisen kampanjan aihetunniste muistuttaa hieman #TurvallinenOulu -aihetunnistetta. Molemmat ovat lähtöisin julkisesta hankkeesta, jossa #TurvallinenOulu oli paljon #kissmyturku -aihetunnistetta viraalisempi.

En pysty perustelemaan aihetunnisteiden valintaa tieteellisiin artikkeleihin nojautuen tutkimuksen eksploratiivisen luonteen vuoksi. Valinnat tehtiin kokeilumielessä ja ne perustuvat lähinnä omaan kiinnostukseen ja vapaaseen päättelyyn.

Sosiaalisen median sisältö voidaan luetella kulttuurituotteeksi, jolloin tutkimusmenetelmänä voidaan käyttää sisällönanalyysia (Uusitalo 1991, 96-97). Tutkimusmenetelmä on tarkalleen ottaen sisällön erittely. Sisällön erittely viittaa lähinnä aineiston määrälliseen kuvailuun, kun taas sisällönanalyysilla viitataan tekstin sisällön sanalliseen kuvailuun (Puusniekka & Saaranen-Kauppinen 2006). Sanojen (esim. paikkakuntien) esiintyvyyden laskeminen dokumentista kuuluu sisällön erittelyyn.

Aineisto kvantifioidaan, eli laadullinen aineisto on muutettu määrälliseen muotoon. Tutkielmassa selvitetään, kuinka monta kertaa jokin Suomen paikkakunnista mainitaan näytteissä. Toisena menetelmänä on käytetty data-analytiikasta tuttua tiedonlouhintaa. Tiedonlouhinta suoritettiin käyttäen mukautettua KDD-prosessia. Tulokset on esitetty tilastoina. #UusimaaLockDown -tutkimuskohteen tilastoista luotiin lisäksi karttaesitys sekä piirakkadiagrammi (ks. liite 1 ja 2).

#### 4.2 Aineistonkeruu ja esittely

Aineisto on ladattu Twitteristä NCapturen avulla. Yhteensä hieman yli kymmenen tuhatta twiittia valittiin tutkittavaksi aihetunnistein: #Oulu (1255), #Helsinki (4872),

#Porvoo (71), #TurvallinenOulu (2197), #KissmyTurku (81) ja #UusimaaLockdown (1636). Aineistojen määrä väheni aineiston käsittelyn aikana, sillä aineistosta siivottiin pois twiitit, joiden metatiedoista ei löytynyt paikkatietoa.

Aihetunniste on valittu ilmaisemaan viestin paikkasidonnaisuutta. Aihetunnisteet asettavat mielestäni viestille teeman sujuvasti ja tehokkaasti. Twiittejä olisi voitu ladata myös hakusanoilla, mutta yksittäiset sanat voivat olla liian irrallisia. Yksittäiset sanat voivat aiheuttaa aineiston käsittelyn aikana epämääräisyyttä ja väärinymmärryksiä, minkä vuoksi valitsin aihetunnisteen aineiston latauskriteeriksi.

Twitter määrittelee, miten paljon annetaan ladata aineistoa (QSR International 2020a). Tämä tekijä on hyvä ottaa huomioon. Ei voi varmuudella sanoa mitä twiittejä Twitter ”antaa” aineistoksi ja mitä ei. Omien havaintojeni perusteella Twitter antaa ladata uusimmat twiitit tiettyyn pisteeseen asti ja tämä piste on jollain tavalla sidottu aikaan. Tämän vuoksi aineistosta voidaan puhua näytteinä otoksien sijaan. Tutkielmassa käytetystä aineistoista voidaan käyttää myös termiä ”datasetti”, mutta tässä tutkielmassa käytän sanoja ”aineisto” ja ”näyte” yksinkertaisuuden nimissä.

Paikkatietoihin viittaavien twiittien jakamiseen osallistuneiden paikkatiedot saadaan viestien metatiedoista kohdasta: location. Location on Twitterin käyttäjän ilmoittama (vapaasti kirjoitettu) sijainti. Location -kohdassa voi käytännössä lukea mitä tahansa. Sama henkilö on voinut siis lähettää useita aineiston twiittejä. Tässä tutkimuksessa on pyrittävä luottamaan siihen, että mahdollisemman moni käyttäjä on syöttänyt Twitteriin oikean paikkatiedon.

Näytteet on jaettu kolmeen kontekstiltaan samankaltaiseen ryhmään:

Ryhmässä A on pelkästään paikkakuntaan viittaavia aihetunnisteita: #Oulu, #Porvoo ja #Helsinki.

Ryhmässä B on paikkakuntien omien hankkeiden aihetunnisteita: #KissmyTurku ja poliittista keskustelua herättänyt #TurvallinenOulu

Ryhmässä C tarkastelun kohteena on yksittäinen aihetunniste, joka liittyy Uudenmaan sulkemiseen: #UusimaaLockdown



#TuvallinenOulu oli alun perin Oulun kaupungin aloittama hanke, jonka päämääränä on lähinnä ehkäistä tai vähentää seksuaalirikoksien määrää (Oulun kaupunki 2019). Aihetunniste loi kiivaan poliittisen ja maahanmuuttajavastaisen keskustelun Twitterissä, jonka vuoksi aihetunnistetta voidaan pitää mielestäni viraalina. Viraalisuus ja viestin poliittinen luonne voi vaikuttaa tutkimustuloksiin, minkä vuoksi pidin aihetunnisteen valintaa tutkimukseen positiivisena asiana.

Ryhmään jakamisen on tarkoitus helpottaa tutkimustulosten esittämistä. Samankaltaiset kontekstit ovat mielestäni hyvä tutkia erikseen, sillä viestien eri kontekstit voivat antaa erilaisia tuloksia. Samankaltaisia viestejä on helpompi verrata keskenään.

#### 4.3 Tutkimuksen toteutus

Tutkimusmenetelmänä toimii riisuttu versio KDD-prosessista. Näytteistä eriteltiin pois kaikki tutkielman kannalta epäoleelliset metatiedot. Jäljelle jäi siis ainoastaan paikkatiedot. Näytteet muutettiin .xls -tiedostoksi, jotta dataa pystyttiin analysoimaan Excelissä ja ArcMapissa. Paikkatiedot olivat sarakkeessa tyyppilliseen listamaiseen tapaan. Jokainen paikkakunta oli omassa solussaan.

Jokaisen ryhmän aineistoja käsitellään ja tutkitaan erikseen samoilla säännöillä. A- ja B-ryhmän näytteet lajiteltiin paikkatietojen (location) mukaan ja vietiin (export) .xls -tiedostoksi. Loin Exceliin yksinkertaisen taulukon, joka käy läpi A-riville olevat solut ja laskee kuinka monesti jokin 310 suomalaisesta paikkakunnasta on mainittu joko suomeksi tai ruotsiksi. Taulukko laskee vain ensimmäiseksi mainitun paikkakunnan. Taulukko ei laske mukaan esimerkiksi kirjoitusvirheitä, taivutusmuotoja, kuntia kuntien sisällä, kaupunginosia eikä myöskään paikkakuntien slangiversioita. Aineistojen valtavan määrän takia en korjannut yhtäkään ryhmän A ja B näytettä manuaalisesti, eli taulukko laskee vain suomen- ja ruotsinkieliset kunnat, jotka esiintyivät oikeinkirjoitettuna.

Kolmannen ryhmän (Ryhmä C) aineisto käsiteltiin tarkemmin; aineistoa siivottiin ja paikkatietoja korjattiin. Aineistoon jätettiin vain suomalaisia paikkakuntia. Maiden,

Läänien, maakuntien, ulkomaalaiset yms. paikkatiedot poistettiin. Paikkatiedoiksi haluttiin vain suomalaisia paikkakuntia. Aineistossa oli jonkin verran paikkatietoja, joita ei ollut kirjoitettu oikein tai olivat muuten epämääräisiä. Eri kieliset paikkakunnat esim. Uleåborg käännettiin Ouluksi, epämääräisesti tai virheellisesti ilmaistut paikkakunnat esim. Hel-sin-ki ja Stadi muutettiin Helsingiksi. Kyseisen aineiston paikkatietojen kappalemäärä tippui siivouksen aikana 1636 kappaleesta 803. Tapaukset laskettiin paikkakunnittain, jonka jälkeen paikkakuntien summat yhdistettiin yksitellen paikkakunnan omiin maakuntiin. Tämän jälkeen tiedot lisättiin Excel-taulukoon, johon myös lisättiin maakuntakohtaiset prosenttiosuudet. Excelin avulla luotiin taulukko (ks. taulukko 6) ja ArcMapin avulla tulos visualisoitiin (ks. liite 1. kartta).

#### 4.4 Tutkielmassa käytetyt ohjelmistot

Käyn tässä kappaleessa lyhyesti läpi tutkielmassa käytettyjä ohjelmistoja. Ohjelmistot, sekä niiden versiot, antavat kuvan tutkimuksen laadukkuudesta. Tietty ohjelmistot ja eri versiot toimivat toisiinsa ohjelmiin ja versioihin nähden eri tavalla. Jotkin ohjelmat ovat puutteellisia ja taas joidenkin ohjelmien versioiden toimivuus vaihtelee, mikä taas vaikuttaa tuloksien laadukkuuteen. Tämän osion on tarkoitus auttaa lukijoita tutkimuksen kriittiseen tarkasteluun eri ohjelmien ja niiden versioiden perusteella. Ohjelmien mainitseminen voi myös auttaa tiedonlouhinnasta kiinnostuneita lataamaan ja kokeilemaan samoja ohjelmia.

##### 4.4.1 NCapture 1.0.290.0

Ncapture on QSR tuottama ilmainen liitännäinen, jota käytetään aineiston lataamiseen. Liitännäinen asennetaan nettiselaimelle esim. Google Chromeen. NCapturen avulla voidaan ladata aineistoa eri sosiaalisista medioista mm. Youtubesta, Facebookista ja Twitteristä. (QSR International 2020b.)

#### 4.4.2 NVivo 12

NVivo on laajasti käytetty QSR Internationalin tuottama maksullinen ohjelmisto, jolla järjestetään, analysoidaan ja visualisoidaan laadullista dataa (McNiff 2016). Tässä tutkielmassa NCapturella kerätty aineisto tuodaan NVivoon, jossa sitä käsitellään. NVivon avulla dataa voidaan esimerkiksi siivota ja muuttaa kunnes datasta voidaan muodostaa yhteyksiä, joita ei aiemmin nähty. NVivon avulla voidaan myös tutkia massadataa, minkä vuoksi ohjelmisto sopii hyvin tämän tutkielman tekoon.

#### 4.4.3 ArcMap 10.7.1

ArcMap on maksullinen paikkatietojärjestelmä, jonka avulla voidaan mm. tuottaa, muokata, visualisoida paikkatietoja. Tässä tutkielmassa käyttäjien paikkatietoja analysoidaan NVivolla ja Excelillä, kun taas ArcMappia käytetään ryhmän C tuloksen visualisointiin (ks. liite 1. karttaesitys).

#### 4.5 Tutkijan massadata -etiikka ja lainsäädäntö

Kuten kaikissa tutkimuksissa, tutkimusetiikka on hyvä ottaa huomioon. Olen pyrkinyt suojaamaan aineistosta saatujen henkilöiden identiteetin, viestit ja muut tiedot mm. tallentamalla tiedot paikalliselle yksityisomistuksessa olevalle kovalevylle. Aineisto myös käsitellään yhdellä ja samalla tietokoneella koko tutkielman teon aikana. Tutkimusdataa ei myöskään siirretä laitteesta toiselle. Tutkimusaineisto tullaan myös tuhoamaan asianmukaisin keinoin. Tutkielmassa ei käy ilmi yksittäisten henkilöiden viestejä tai muitakaan tietoja, joiden avulla henkilö voidaan tunnistaa. Twiitit ovat julkisesti kaikille näkyviä, ellei niitä asetuksista määritellä toisin. Tässä tutkielmassa ei ole käytetty yksityisiä tai piilotettuja twiittejä (aineiston lataamisen aikana).

#### 4.5.1 Massadataan liittyvä tutkimusetiikka

Sosiaalisen massadatan käyttö on lisääntynyt erilaisissa tutkimuksissa ja siihen liittyvä tutkimusetiikka sekä sen määrittely on osoittautunut hankalaksi aiheeksi (Kosonen 2018, 117). Fieslerin ja Proferesin (2018, 8) tutkimuksen mukaan osa heidän tutkimukseensa osallistuneista oli huolissaan siitä, että heidän twiittinsa voitaisiin yhdistää heidät heidän todellisiin identiteetteihinsä. On tärkeää pitää huoli siitä, että dataa käsitellään oikeaoppisesti, jottei yksilöiden oikeuksia ja yksityisyyttä loukata. Sosiaalisessa massadataan liittyvä tutkimusetiikka voi olla usealle vaikea määritelmä. Samaan aikaan sosiaalista mediaa tutkiva voi haluta mahdollisemman kattavan ja sensuroimattoman aineiston kuin mahdollista, mutta yksilön oikeudet ovat myös yhtä tärkeitä. Toisaalta kun käsitellään massadataa niin yhden henkilön data hukkuu helposti valtavaan tietomäärään, joten sinällään kenenkään ei tarvitse huolehtia liikaa asiasta.

#### 4.5.2 Lainsäädäntö

Euroopan parlamentin ja neuvoston asetus (2016/679) määrittää henkilötietojen käsittelyn tieteelliseen tarkoitukseen laillisena käsittelytoimena kohdassa 50. Henkilötieto on tietoa, jonka avulla henkilö voidaan tunnistaa esim. tietoja yhdistämällä (Tietosuojavaltuutetun toimisto 2020).

Kunnioittaakseni ihmisten oikeuksia, olen datan siivoamisvaiheessa siivonnut pois mahdollisemman paljon metatietoja, jotka eivät ole tutkimuksen kannalta välttämättömiä. Tämä mahdollistaa sen, että datan siivouksen jälkeen yksilöitä on vaikeampi tunnistaa esim. pelkkien paikkatietojen avulla. Poikkeuksia voi tietenkin olla, sillä en ole manuaalisesti käynyt läpi jokaista kerättyä twiittiä, ja sen metatietoja, johtuen datan massiivisesta määrästä.

#### 4.5.3 Lisenssit ja luvat

Twitter (2020c) sallii Twitteristä saatujen aineistojen käyttämisen ei-kaupalliseen tarkoitukseen. Liite 2. kartta-aineistona toimii Maanmittauslaitoksen aineisto: ”hallintorajat, teemakartoille, ei merialueita”, 2020 (CC 4.0 -lisenssi).

## 5 TULOKSET

Tässä kappaleessa esittelen valitsemani aihetunnisteet ja niiden tulokset. Tulokset on esitetty taulukkomuodossa. Kappaleet etenevät ryhmittäin (A-C). Kirjain 'n' viittaa twiittien määrää, suluissa oleva luku viittaa prosenttiosuuteen. Taulukkojen tulokset on järjestetty paikkakunnittain suurimmasta pienempään, liittyen twiittien määrästä. Poikkeuksena ryhmän C taulukko, joka on järjestetty maakunnittain. Taulukkojen vasempaan alakulmaan on ilmoitettu tuntemattomien paikkatietojen määrä. Tuntemattomat paikkatiedot ovat niitä tietoja, joita aikaisemmin koodattu taulukko ei pystynyt tunnistamaan. Tuntemattomat paikkatiedot ovat siis kaikkea paitsi oikeinkirjoitettuja suomalaisia paikkakuntien nimiä. Näitä tietoja ei laskettu mukaan yhteismäärään. Tulokset on pyöristetty yhden desimaalin tarkkuuteen.

### 5.1 Ryhmän A tulokset

Ryhmän A tulokset koostuivat teemaltaan samankaltaisista näytteistä, keskittyen pelkästään paikkakuntiin viittaaviin aihetunnisteisiin. Näitä aihetunnisteita olivat #Helsinki (ks. taulukko 1), #Oulu (ks. Taulukko 2) ja #Porvoo (ks. Taulukko 3).

Ryhmä A:n tilastoista (ks. taulukot 1-3) näkee, miten paikkaan liittyviä aihetunnisteita, ja niitä twiittaavien käyttäjien fyysinen, sijainti korreloituu paikkaan viittaavien aihetunnisteiden kanssa. Peräti 93,3% #Helsinki -aihetunnistetta twiittanneiden paikkakunnaksi oli ilmoitettu Helsinki (ks. taulukko 1). 85% #Oulu -aihetunnistetta twiittanneiden paikkakunnaksi oli ilmoitettu Oulu (ks. taulukko 2) ja #Porvoo -aihetunnistetta twiittasi lähinnä porvoolaiset (83,8%) (ks. taulukko 3). Näiden näytteiden perusteella voidaan nähdä yhteys paikallisuuden ja paikkaan viittaavan twiitin kanssa – paikalliset viestittelevät aktiivisemmin omaan paikkakuntaan liittyvistä twiiteistä. Tässä tutkielmassa ei tehty syvällistä sisällönanalyysia, jonka avulla voitaisiin selvittää missä yhteyksissä esimerkiksi helsinkiläiset twiittasivat #Helsinki -aihetunnisteellisia twiittejä. Sisällönanalyysilla voisi olla mahdollista myös selvittää

millaiset viestit leviävät parhaiten ja miksi. Jatkotutkimuksiin voisi tuoda mukaan sisällönanalyysi ilmiön paremman ymmärtämisen vuoksi.

Taulukko 1. #Helsinki -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain (n=2778).

#Helsinki -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain		Paikkakunnat				
Helsinki	Tampere	Espoo	Lappeenranta	Turku	Muu Suomi	Yhteensä
n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)
2593 (93.3)	66 (2.4)	23 (0.8)	19 (0.7)	17 (0.6)	60 (2.2)	2778 (100)
1039 paikkatiedoista tuntemattomia						

Taulukko 2. #Oulu -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain (n=920).

#Oulu -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain			Paikkakunnat				
Oulu	Helsinki	Espoo	Tampere	Kajaani	Lapua	Muu Suomi	Yhteensä
n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)
782 (85)	128 (13.9)	3 (0.3)	3 (0.3)	2 (0.2)	2 (0.2)	3 (0.3)	920 (100)
218 paikkatiedoista tuntemattomia							

Taulukko 3. #Porvoo -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain (n=37).

#Porvoo -aihetunnistetta twiitanneiden paikkatiedot paikkakunnittain		Paikkakunnat		
Porvoo	Helsinki	Espoo	Muu Suomi	Yhteensä
n (%)	n (%)	n (%)	n (%)	N (%)
31 (83.8)	5 (13.5)	1 (2.7)	0 (0)	37 (100)

16 paikkatiedoista tuntemattomia

#Helsinki -aihetunnisteen tuntemattomat paikkatiedot (1039 kappaletta) (ks. taulukko 3) olivat suhteutettuna korkeat muihin ryhmän aihetunnisteisiin verrattuna lähinnä siksi, että moni paikkatiedoltaan ulkomaalainen twiittasi kyseistä aihetunnistetta enemmän verrattuna muihin ryhmän tapauksiin. Äskeinen väite perustuu omakohtaiseen havainnointiin. On hyvä huomioida, että #Porvoo -aihetunnisteen kohdalla näyte oli hyvin pienikokoinen (37 kappaletta). Näytteen koko vaikuttaa tulosten tarkkuuteen; mitä suurempi näyte, sitä luotettavampi tulos.

## 5.2 Ryhmän B tulokset

Ryhmän B tulokset (ks. taulukot 4 ja 5) muistuttavat hyvin paljon ryhmän A tuloksia. Ryhmät olivat kontekstiltään eriäviä; Ryhmä A koostui aihetunnisteista, joissa mainittiin pelkästään paikkakunta, kun taas ryhmä B koostui aihetunnisteista, joissa käy ilmi paikkakunta sekä hankkeen nimi. #kissmyturku -aihetunnisteen twiittaajat olivat lähinnä Twitter-käyttäjiä, jotka ilmoittivat paikkakunnakseen Turku, peräti 97,6% (ks. taulukko 4). Tässäkin tapauksessa on hyvä ottaa huomioon näytteen pieni koko (42 kappaletta). Taulukkoon valittiin esitettäväksi kaksi suurinta paikkakuntaa, muu Suomi sekä tuntemattomien paikkatietojen määrä.

#TurvallinenOulu -tapauksen kohdalla tulokset olivat samansuuntaiset aikaisempien tulosten kanssa, eli 55,3% kyseisen aihetunnisteen twiittaajista ilmoittivat paikkakunnakseen Oulu (ks. taulukko 5). Aihetunnisteen viraalisuus saattaa olla syy



alhaiseen paikallisten määrään verrattuna muihin tuloksiin. Keskustelun kiivas, poliittinen ja aggressiivinen luonne saattoi myös vaikuttaa siihen, miksi tuntemattomien paikkatietojen määrä oli niin korkea. Paikkatiedoltaan tuntemattomien määrä oli peräti 1009 kappaletta, kun taas 903 twiittien lähettäjien paikkatiedot voitiin tunnistaa (ks. taulukko 5). Twiittaajista 830 oli ilmoittanut paikkatiedokseen pelkän Suomen. Paikkatieto oli siis tässä tapauksessa muiden aineistojen paikkatietoihin verrattuna melko epätasmallista. Kyseiseen taulukkoon valittiin paikkatiedoltaan viisi suurinta paikkakuntaa, muu Suomi sekä tuntemattomien paikkatietojen määrä.

Taulukko 4. #kissmyturku -aihetunnistetta twiittanneiden paikkatiedot paikkakunnittain (n=42).

#kissmyturku - aihetunnistetta twiittanneiden paikkatiedot paikkakunnittain	Paikkakunnat			
	Turku	Helsinki	Muu Suomi	Yhteensä
	n (%)	n (%)	n (%)	N (%)
	41 (97.6)	1 (2.4)	0 (0)	42 (100)

11 paikkatiedoista tuntemattomia

Taulukko 5. #TurvallinenOulu -aihetunnistetta twiittanneiden paikkatiedot paikkakunnittain (n=903).

#TurvallinenOulu - aihetunnistetta twiittanneiden paikkatiedot paikkakunnittain	Paikkakunnat					
	Oulu	Helsinki	Tampere	Vantaa	Joensuu	Muu Suomi
	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)
	499 (55.3)	167 (18.5)	99 (11)	54 (6)	34 (3.8)	50 (5.5)
						903 (100)

1009 paikkatiedoista tuntemattomia

Halu olla anonyymi, saattaa olla epämääräisten paikkatietojen takana. Tutkielmassani käytiin aiemmin läpi anonyymiuden yhteyttä aggressiiviseen verkkokäyttäytymiseen. Moni (paikkatiedoltaan) anonyymi käyttäjä oli osallistunut keskusteluun ja luultavasti anonyymien käyttäjien määrän vuoksi keskustelu oli ajoittain kiihkeää, jopa aggressiivista. Edelleenkin, tässä tutkimuksessa ei tutkita sisältöä, vaikka sisällönkuvaus toisi varmasti esiin kiinnostavia seikkoja.

### 5.3 Ryhmän C tulokset

Ryhmän C kohdalla tarkastuksen kohteena on twiitit liittyen aihetunnisteseen #UusimaaLockDown. Taulukkoon valittiin maakunniksi Uusimaa, sen naapurimaakunnat sekä Pirkanmaa kuin myös muu Suomi. Ensiksi käydään läpi Uusimaa ja sen naapurimaakunnat. Tuntemattomia paikkatietoja ei ole, sillä tiedot käsiteltiin manuaalisesti ja virheelliset tiedot korjattiin tai poistettiin, samoitten myös Suomen ulkopuoliset paikat poistettiin.

Ryhmän C tulokset myötäilevät jonkin verran aikaisempien ryhmien tuloksia. Siivotusta datasta 76.3% #UusimaaLockDown -aihetunnistetta twiitanneiden käyttäjien paikkatiedot pystyttiin paikantamaan Uudenmaan sisälle (ks. taulukko 6). Uudenmaan naapurimaakunnat (Päijät-Häme, Kanta-Häme, Varsinais-Suomi ja Kymmenlaakso) muodostivat yhteensä 83,8% Twiiteistä (673/803).

Taulukko 6. #UusimaaLockDown -aihetunnistetta twiitanneiden paikkatiedot maakunnittain (n=803).

#UusimaaLockDown - aihetunnistetta twiitanneiden paikkatiedot maakunnittain							Maakunnat	
Uusimaa	Päijät- Häme	Kanta- Häme	Varsinais- Suomi	Kymmen- laakso	Pirkanmaa	Muu Suomi	yhteensä	
n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)	
613 (76.3)	15 (1.9)	15 (1.9)	25 (3.1)	5 (0.6)	69 (8.6)	61 (7.6)	803 (100)	

Syy siihen, miksi Pirkanmaa oli toiseksi suurin maakunta, jää vielä avoimeksi. Kenties pirkanmaalaisia matkustaa Uudenmaan alueelle paljon, jonka vuoksi pirkanmaalaiset viestittelivät kyseistä aihetunnistetta toiseksi eniten – Uudenmaan sulku ehkä vaikutti heidän elämäänsä niin paljon, että se motivoi pirkanmaalaisia twiittaamaan aiheesta. Kyseisen asian selvittämiseksi vaadittaisiin jo vähän työläämpiä tutkimuksia. Kyselyitä pitäisi lähettää twiittaajille ja kyselyiden kysymysasetelman tulisi olla tarpeeksi hyvin toteutettu, jotta siitä saataisiin selville twiittaajien motivaatio kyseisen aihetunnisteen levittämiseen. Twitterin omat säännöt voivat tosin kieltää kyselytutkimuksen teettämistä heidän käyttäjilleen. Jatkotutkimuksia ajatellen olisi hyvä ottaa selvää minkälaisia tutkimuksia käyttäjistä saa, tai ei saa, toteuttaa.

## 6 POHDINTA

Tutkielmassa tarkasteltiin paikkaan viittaavan tiedon leviämistä sosiaalisessa mediassa tiedon levittäjien paikkatietoon nähden. Sosiaalisen median tutkimukset ja informaatiokäyttäytyminen kuuluvat informaatiotutkimuksen tutkimuskenttiin. Tutkielman aihetta käytiin läpi monitieteellisesti, mikä on tyypillistä informaatiotutkimuksessa. Tässä kappaleessa käydään vielä hieman läpi tutkimusta ja tutkimukseen vaikuttavia tekijöitä. Myös tuloksia käydään syvemmin läpi sekä pohditaan erilaisia jatkotutkimusmahdollisuuksia.

Tutkimuksen tulosten perusteella paikalliset levittävät omaan paikkakuntaansa liittyviä viestejä Twitterissä aktiivisemmin ulkopaikkakuntalaisiin verrattuna. Sosiaalisen mediassa luodut käyttäjien väliset suhteet muistuttavat enemmän, tai vähemmän, fyysisen maailman ystävyssuhteita (Currie ym. 2012, 23). Luulen, että paikallinen ystäväpiiri on osittain vastuussa ilmiön selittämisen kannalta. Olen myös sitä mieltä, että ihmisille paikalliset tapahtumat ja asiat itsessään herättävät enemmän tunteita. Tunteet taas hallitsevat ihmisen (informaatio)käyttäytymistä (Savolainen 2015, 1; Nahl 2006; Dang-Xuan & Stieglitz 2013, 222-224).

Uskoisin, että tutkielmassa aiemmin mainittu self-reference effect vaikutti ilmiöön. Kyseinen ilmiö tarkoitti sitä, että ihminen käsittelee informaatiota tarkemmin, jos se liittyy häneen itseensä. Esimerkiksi #UusimaaLockDown -aihetunnistetta twiitattiin lähinnä Uudenmaan alueella, koska kyseinen tilanne vaikutti siellä asuvien ihmisten elämään. Uskon, että tällä ilmiö, tai sen kaltainen ilmiö, voisi selittää sen miksi paikalliset levittävät ulkopaikkakuntalaisia aktiivisemmin omaan paikkakuntaan liittyviä twiittejä.

Edellä mainitut tekijät ovat mielestäni mahdollisia selityksiä sille, miksi suomalaiset twiittaavat paikallisia twiittejä ulkopaikkakuntalaisia aktiivisemmin. Myös paikkaan liittyvät viestit, eli viestien kontekstit, voivat vaikuttaa viestien leviämiseen. #TurvallinenOulu oli kontekstiltaan poliittisesti varautunut aihetunniste ja se levisi viraalisemmin kuin mikään muu tutkimuksen aihetunniste. #TurvallinenOulu -aihetunnisteen levittäjien motiivina saattoi myös toimia aikaisemmin mainittu vahvistusharha. Kyseisen aihetunnisteen levittäjät joko puolsivat Turvallinen Oulu -

hanketta tai olivat sitä vastaan. Paikallisuuden tärkeydestä viestien leviämiseen suhteen on ollut näyttöä; Palenin ja Starbirdin tutkimus osoitti (2012) miten Kairolaiset twiittasivat aktiivisesti Egyptin vallankumouksen aikana 2011. Syitä sille, miksi paikkaan sidotut leviävät niin kuin ne leviävät, on varmasti monia. Syiden selvittelyyn vaaditaan erilaisia tutkimusmenetelmiä ja mahdollisesti myös erilaisia aineistonhankintamenetelmiä.

Tässä tutkielmassa ei keskitytty siihen, miksi paikallinen tieto leviää Twitterissä. Tutkimuksessa pyrittiin löytämään vastaus siihen, miten paikkaan sidottu tieto leviää ja miten suuri vaikutus käyttäjän fyysisellä sijainnilla on. Mielestäni tutkimus vastasi molempiin kysymyksiin. Ensimmäinen kysymys, miten paikkaan sidottu tieto leviää, sai vastauksensa. Paikkaan sidotut twiitit leviävät pääasiassa paikallisten toimesta. Toiseen kysymykseen, miten suuri vaikutus käyttäjän fyysisellä sijainnilla on tiedon leviämiseen kannalta, on hankalempi vastata. Fyysisen sijainnin merkittävyyttä on hankalempi määritellä. Näytteistä ”heikoin” tapaus (#TurvallinenOulu), koostui 55,3% oululaisista twiittaajista (taulukko 5). Oulun väkiluku on vähän päälle 200 000 (Oulun Kaupunki 2019). Suomen väkiluku on n. 5,5 miljoonaa (Tilastokeskus 2020). 205 000 jaettuna 5,5 miljoonalla on n. 3,7%. Pitäisin itse fyysistä sijaintia merkittävänä, jos noin 3,7% väestöstä oli vastuussa 55,3% tietyn näytteen twiiteistä. Tutkimuksen eksploratiivisen luonteen vuoksi, tutkimuksessa sinällään ei kumottu, vahvistettu tai luotu uusia teorioita. Tutkielman tarkoitus oli kartoittaa ennestään tutkimatonta tutkimuskohdetta, eli suomalaisten Twitter-käyttäjien tapaa levittää paikkaan viittaavia twiittejä. Tutkielma tarjoaa kuitenkin päätuloksen: paikkakunta-aiheisia twiittejä twiittaa lähinnä paikalliset. Tutkielma myös toivottavasti tarjoaa jonkinlaisen alustan muille, jotka haluavat tutkia samaa, tai samankaltaista, ilmiötä. Tutkimuskohde oli ennestään tuntematon, joten toivottavasti tämä tutkielma teki siitä vähemmän tuntemattoman.

Vaikka tutkielman tulokset myötäilevät toinen toistaan, niin on otettava huomioon muutamia tutkimuksessa esille nousseita seikkoja. Twitterin algoritmit määrittelevät sen miten paljon aineistoa voidaan ladata kerralla. Tämä voidaan korjata esimerkiksi parilla tavalla: voidaan ladata aineistoa kuukausien tai vuosien aikana useamman kerran. Omakohtaisiin kokemuksiin perustuen, Twitter antaa ladata aineistoa muutaman viikon

päähän aina kerralla. Tietyn kestoisen aineistonkeruun jälkeen, aineistot voidaan yhdistää tutkimuksen aikana isommaksi kokonaisuudeksi. Toinen tapa on ottaa yhteyttä Twitteriin ja sopia erikseen heidän kanssaan datankäytöstä, jolloin voidaan saada ladattua isompia aineistoja.

Jotkin aineistot olivat kooltaan hyvinkin pieniä, etenkin #Porvoo (37) ja #kissmyturku (42) (ks. taulukot 3 ja 4). Jatkotutkimuksissa olisikin suotavaa tutkia mahdollisemman isoja datamääriä luotettavan tuloksen aikaansaamiseksi. Yksi erittäin aktiivinen twiittäjä (esim. super spreader), tai julkisuuden henkilö (influential spreader), voi muuttaa tuloksia suuntaan tai toiseen. Jos toteuttaisin tulevaisuudessa samankaltaisen tutkimuksen niin laskisin paikkatiedot henkilöittäin, en twiiteittäin. Jokainen käyttäjä voitaisiin siis laskea vain kerran. Tässä tutkielmassa sama käyttäjä on voitu laskea mukaan monta kertaa, jos hän on twiitannut useamman kerran käyttäen jotain tutkimuksien aiheutunnisteista. Toisin sanoen, haluaisin siis poissulkea aktiiviset twiittäajat ja tarkastella tiedon leviämistä ihmiseltä toiselle pikemminkin kuin twiittääjien aktiivisuutta. Myös eri ohjelmien kokeilu ja vertaaminen voisi olla tuottavaa. Toisin mukaan myös sisällönanalyysin, jotta voitaisiin saada parempi käsitys siitä, miksi jotkin twiitit leviävät paremmin mitä toiset.

Toivon, että tutkielmasta on apua verkkoviestinnästä, sosiaalisesta massadatasta tai tiedonlouhinnasta kiinnostuneille. Tutkielman tuloksista voisi olla apua heille, jotka haluavat tutkia aiheen ilmiötä syvemmin. Aiheen jatkotutkimuksista voisi olla apua vaikkapa kunnille, jotka haluavat lisätä näkyvyyttä sosiaalisessa mediassa. Myös turismiin liittyvät yritykset ja julkiset tahot, jotka haluavat markkinoida paikkakuntiaan sosiaalisessa mediassa, voisivat olla kiinnostuneita jatkotutkimuksista.

## 7 LÄHTEET

Ahola, S., Bäckström, A. & Mäkinen, J.-P. (2014). Ruokariskit ja -asiantuntijat ruokaskandaalien aikakaudella. Teoksessa: Ahokas, M., Ahola, S., Myyry L. & Sakki, I. (toim.), *Arki ajattelu, tieto ja oikeudenmukaisuus* (s.113-129). Sosiaalitieteiden laitoksen julkaisuja 2014:18. Helsinki: Helsingin yliopisto

Anttila, J., Jokinen, E. & Ruonen, T. (2018). Vahvistusharhan tunnistaminen ja ehkäiseminen johtajan päätöksenteossa. Teoksessa: Hytönen, K., Matvejeff, P. & Suomala, J (toim.), *Päätöksenteon ilmiöt johtajan arjessa* (s.42-53). Laurea julkaisut. Helsinki: Laurea-ammattikorkeakoulu.  
<https://www.theseus.fi/bitstream/handle/10024/146539/Laurea%20julkaisut%2093.pdf?sequence=1&isAllowed=y#page=22>

Barak, A. & Lapidot-Lefer, N. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*. 28(2), 434-443. DOI: 10.1016/j.chb.2011.10.014

Barbera, P., Guess, A., Nyhan, B., Sanovich, S., Siegel, A., Stukal, D., Tucker, J. & Vaccari, C. (2018). Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. DOI: 10.2139/ssrn.3144139

Barlett, C. (2015). Anonymously hurting others online: The effect of anonymity on cyberbullying frequency. *Psychology of Popular Media Culture* 4(2), 70-79. DOI: 10.1037/a0034335.

Bik, H. & Goldstein, M. (2013). An Introduction to Social Media for Scientists. *PLoS Biology* 11(4), e1001535. DOI: 10.1371/journal.pbio.1001535

Bramer, M. (2013). Introduction to Data Mining. *Principles of Data Mining, Undergraduate Topics in Computer Science*. DOI: 10.1007/978-1-4471-4884-51

Bruce, H., Fidel, R. & Pettigrew, K. (2001). Conceptual Frameworks in Information Behavior. Annual Review of Information Science and Technology (ARIST) 35, 43-77. <http://faculty.washington.edu/fidelr/RayaPubs/ConceptualFrameworks.pdf> (käytetty 5.4.2020).

Chatman, E. A. (1991). Life in a Small World: Applicability of Gratification Theory to Information-seeking Behavior. *Journal of the American Society for Information Science (1986-1998)*, 42(6), 438.  
<https://search.proquest.com/docview/216894783?accountid=13031> (käytetty 5.4.2020).

Chatman, E. A. (1996). The Impoverished Life-World of Outsiders. *Journal of the American Society for Information Science (1986-1998)*, 47(3), 193.  
<https://search.proquest.com/docview/216900346?accountid=13031> (käytetty 5.4.2020).

Currie, D., Jutemar, E., Kaltenbrunner, A., Laniado, D., Mascolo, C., Scellato, S. & Volkovich, Y. (2012). Far from the eyes, close on the web: Impact of geographic distance on online social interactions. DOI: 10.1145/2342549.2342555

Dang-Xuan, L. & Stieglitz, S. (2013). Emotions and Information Diffusion in Social Media — Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems* 29(4), 217-248. DOI: 10.2753/MIS0742-1222290408

Davis, C., Ferrara, E., Flammini, A., Menczer, F. & Varol, O. (2017). Online Human-Bot Interactions: Detection, Estimation, and Characterization. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, 280-289. <https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15587/14817>

Dewing, M. (2012). Social Media: An Introduction. *Publication No. 2010-03-E*. <https://bdp.parl.ca/staticfiles/PublicWebsite/Home/ResearchPublications/InBriefs/PDF/2010-03-e.pdf> (käytetty 27.3.2020).

Dunbar, R., Mesoudi, A. & Whiten, A. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology* 97(3), 405-423. <http://search.ebscohost.com.pc124152.oulu.fi:8080/login.aspx?direct=true&db=c8h&AN=106224210&site=ehost-live> (käytetty 5.4.2020).

*Euroopan parlamentin ja neuvoston asetetus (EU) 2016/679*. EUR-Lex. <https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX%3A32016R0679>

Faust, K. & Wasserman, S. (1994). *Social Network Analysis. Methods and Applications*. USA: Cambridge University Press 1994.

Favre, C., Guille, A., Hacid, H. & Zighed, D. (2013). Information Diffusion in Online Social Networks: A Survey. *ACM SIGMOD Record* 42(2), 17-28. DOI: 10.1145/2503792.2503797

Fiesler, C. & Proferes, N. (2018). "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4(1), 1-14. DOI: 10.1177/2056305118763366

Funder, D. & Krueger, J. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences* 27(3). 313-27; discussion 328. DOI: 10.1017/S0140525X04000081

Haasio, A. (2015). Disinformatiivinen ja normatiivinen informaatio. *Informaatiotutkimus* 34 (4), 2015. <https://journal.fi/inf/article/view/53512/16669> (käytetty 5.4.2020)

Han, J., Kamber, M. & Pei, J. (2012). *Data Mining. Concepts and Techniques*. USA: Elsevier.

Hatano, N. & Kawamoto, T. (2014). Viral spreading of daily information in online social networks. *Physica A: Statistical Mechanics and its Applications* 406, 34-41. DOI: 10.1016/j.physa.2014.03.054



- Hermkens, K., Kietzmann, J., McCarthy, I. & Silvestre, B. (2011). Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media. *Business Horizons* 54(3), 241-251. DOI: 10.1016/j.bushor.2011.01.005
- Heinonen, S. (2009). Sosiaalinen media. Avauksia nettiyhteisön maailmaan ja vuorovaikutuksen uusiin muotoihin. *TUTU-eJULKAISUJA* 1/2009. URL: [https://www.utupub.fi/bitstream/handle/10024/147581/eTutu\\_2009-1.pdf?sequence=1](https://www.utupub.fi/bitstream/handle/10024/147581/eTutu_2009-1.pdf?sequence=1) (käytetty 5.4.2020).
- Hoang, T-A, Kywe, S., Lim, E-P. & Zhu, F. (2012). On Recommending Hashtags in Twitter Networks. *SocInfo 2012: Social Informatics* 7710, 337-350. DOI: 10.1007/978-3-642-35386-4\_25
- Johnson, B. & Symons, C. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, Vol 121(3), 371-394. DOI: 10.1037/0033-2909.121.3.371
- Kawamoto, T. (2013). A stochastic model of tweet diffusion on the Twitter network. *Physica A: Statistical Mechanics and its Applications* 392(16), 3470-3475. DOI: 10.1016/j.physa.2013.03.048
- Kihlstrom, J. & Klein, S. (1986). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General*, Vol 115(1), 26-38. DOI: 10.1037/0096-3445.115.1.26
- Kolmes, S. (2011). Climate Change: a Disinformation Campaign. *Environment Science and Policy for Sustainable Development* 53(4), 33-37. DOI: 10.1080/00139157.2011.588553
- Kosonen, M., Laaksonen, S-M., Rydenfelt, H. & Terkamo-Moisio, A. (2018). Sosiaalinen media ja tutkijan etiikka. *Media & viestintä* 41(2018):1, 117–124. <https://journal.fi/mediaviestinta/issue/view/4738> (käytetty 5.4.2020)
- Kotimaisten kielten keskus (2020). Kielitoimiston sanakirja. <https://www.kielitoimistonsanakirja.fi/#/data> (käytetty 27.3.2020).
- Kuhlthau, C. (2020). Information Search Process. <https://wp.comminfo.rutgers.edu/ckuhlthau/information-search-process/> (käytetty 2.4.2020).
- Laney, D. (2001). 3D data Management: Controlling Data Volume, Velocity, and Variety. *ADS* 6 Feb 01.949 Addendum. <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (käytetty 27.3.2020).
- Latora, V., Mascolo, C., Musolesi, M. & Scellato, S. (2010). Distance matters: geo-social metrics for online social networks. *WOSN'10: Proceedings of the 3rd Wonference on Online social networks*, 8-16. [https://static.usenix.org/events/wosn10/tech/full\\_papers/Scellato.pdf](https://static.usenix.org/events/wosn10/tech/full_papers/Scellato.pdf) (käytetty 5.4.2020).

Matikainen, J. (2008). Sosiaalinen media – millaista sosiaalisuutta? *Media & Viestintä*, 31(4). DOI: 10.23983/mv.63013

McNiff, K. (2016). What is Qualitative Research? <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/resources/blog/what-is-qualitative-research> (käytetty 5.4.2020).

Mäkinen, I. (2010). Informaatiotutkimuksen tie. Teoksessa: S. Serola (toim.), *Ote informaatiosta: johdatus informaatiotutkimukseen ja interaktiiviseen mediaan* (s. 11-74). Vantaa: BTJ Finland Oy.

Nahl, D. (2006). Affective and cognitive information behavior: Interaction effects in Internet use. *Proceedings of the American Society for Information Science and Technology* 42(1). DOI: 10.1002/meet.1450420196

Oulun kaupunki (2019). *Tilastoja Oulusta*. <https://www.ouka.fi/oulu/oulu-tietoa/tilastoja-oulusta>

Oulun kaupunki (2020). Turvallinen Oulu. <https://www.ouka.fi/oulu/kehittamishankkeet/turvallinen-oulu-hanke> (käytetty 17.4.2020).

Palen, L. & Starbird, K. (2012). (How) Will the Revolution be Retweeted? Information Diffusion and the 2011 Egyptian Uprising. *CSCW '12 Computer Supported Cooperative Work*. DOI: 10.1145/2145204.2145212

Puusniekka, A. & Saaranen-Kauppinen, A. (2006). Sisällönanalyysi. [https://www.fsd.tuni.fi/menetelmaopetus/kvali/L7\\_3\\_2.html](https://www.fsd.tuni.fi/menetelmaopetus/kvali/L7_3_2.html) (käytetty 17.4.2020).

QSR International (2020a). NCapture. <https://help-nv.qsrinternational.com/12/win/v12.1.90-d3ea61/Content/ncapture/ncapture.htm#MiniTOCBookMark7> (käytetty 5.4.2020).

QSR International (2020b). About NCapture. [http://help-nv11.qsrinternational.com/desktop/concepts/about\\_ncapture.htm](http://help-nv11.qsrinternational.com/desktop/concepts/about_ncapture.htm) (käytetty 5.4.2020).

Saari, M. (2016). *Tavoitteellinen sisältömarkkinointi sosiaalisessa mediassa*. Opinnäytetyö. Helsinki: Metropolian ammattikorkeakoulu

Sairanen, A. & Savolainen, R. (2008). "Mä en halua tietää". Terveystieteen välttämisen syyt ja muodot. *Informaatiotutkimus* 27(3), 1-11. <https://journal.fi/inf/article/view/724/590> (käytetty 5.4.2020).

Salmi, H. (2018). Viraalisuus – kulttuurihistoriallinen näkökulma. *niin & näin* 1/18, 71-79. <https://netn.fi/sites/www.netn.fi/files/netn181-11.pdf> (käytetty 5.4.2020).

Savolainen, R. (2015). The interplay of affective and cognitive factors in information seeking and use: comparing Kuhlthau's and Nahf's models. *Journal of Documentation* 71(1), 175-197. DOI: 10.1108/JD-10-2013-0134

Suomen Standardisoimisliitto (2020). Paikkatieto.  
[https://www.sfs.fi/standardien\\_laadinta/sfs\\_n\\_standardisointiryhmat/it-standardisointi/it\\_-\\_aihealueet/paikkatieto](https://www.sfs.fi/standardien_laadinta/sfs_n_standardisointiryhmat/it-standardisointi/it_-_aihealueet/paikkatieto) (käytetty 29.3.2020).

The World Bank (2020). *Individuals using the Internet (% of population) - Egypt, Arab Rep.* International Telecommunication Union, World Telecommunication/ICT Development Report and database.  
<https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=EG>

Tieteen Termipankki (2020). Homofilia.  
<https://tieteentermipankki.fi/wiki/Sosiaalipsykologia:homofilia> (käytetty 1.4.2020).

Tietosuoja-valtuutetun toimisto (2020). Mikä on henkilötieto? <https://tietosuoja.fi/mika-on-henkilotieto> (käytetty 29.3.2020).

Tilastokeskus (2020). *Suomen ennakkoväkiluku helmikuun lopussa 5 526 774.*  
[https://www.stat.fi/til/vamuu/2020/02/vamuu\\_2020\\_02\\_2020-03-24\\_tie\\_001\\_fi.html](https://www.stat.fi/til/vamuu/2020/02/vamuu_2020_02_2020-03-24_tie_001_fi.html)  
(käytetty 16.4.2020).

Tuomi, J. (2007). *Tutki ja lue: Johdatus tieteellisen tekstin ymmärtämiseen.* Helsinki: Tammi.

Twitter (2020a). How to Tweet. <https://help.twitter.com/en/using-twitter/how-to-tweet> (käytetty 29.3.2020).

Twitter (2020b). Q4 and Fiscal Year 2019 Letter to Shareholders.  
[https://s22.q4cdn.com/826641620/files/doc\\_financials/2019/q4/Q4-2019-Shareholder-Letter.pdf](https://s22.q4cdn.com/826641620/files/doc_financials/2019/q4/Q4-2019-Shareholder-Letter.pdf) (käytetty 29.3.2020).

Twitter (2020c). Developer Agreement and Policy.  
<https://developer.twitter.com/en/developer-terms/agreement-and-policy> (käytetty 29.3.2020).

Uusitalo, H. (1991). *Tiede, tutkimus ja tutkielma: Johdatus tutkielman maailmaan.* Helsinki: WSOY

Valtiovarainministeriö (2008). VAHTI 8/2008 Valtionhallinnon tietoturvasanasto.  
<https://www.vahtiohje.fi/web/guest/maaritelmat-d> (käytetty 9.4.2020).

Wiki.org (2002). What is wiki. <http://www.wiki.org/wiki.cgi?WhatIsWiki> (käytetty 27.3.2020).

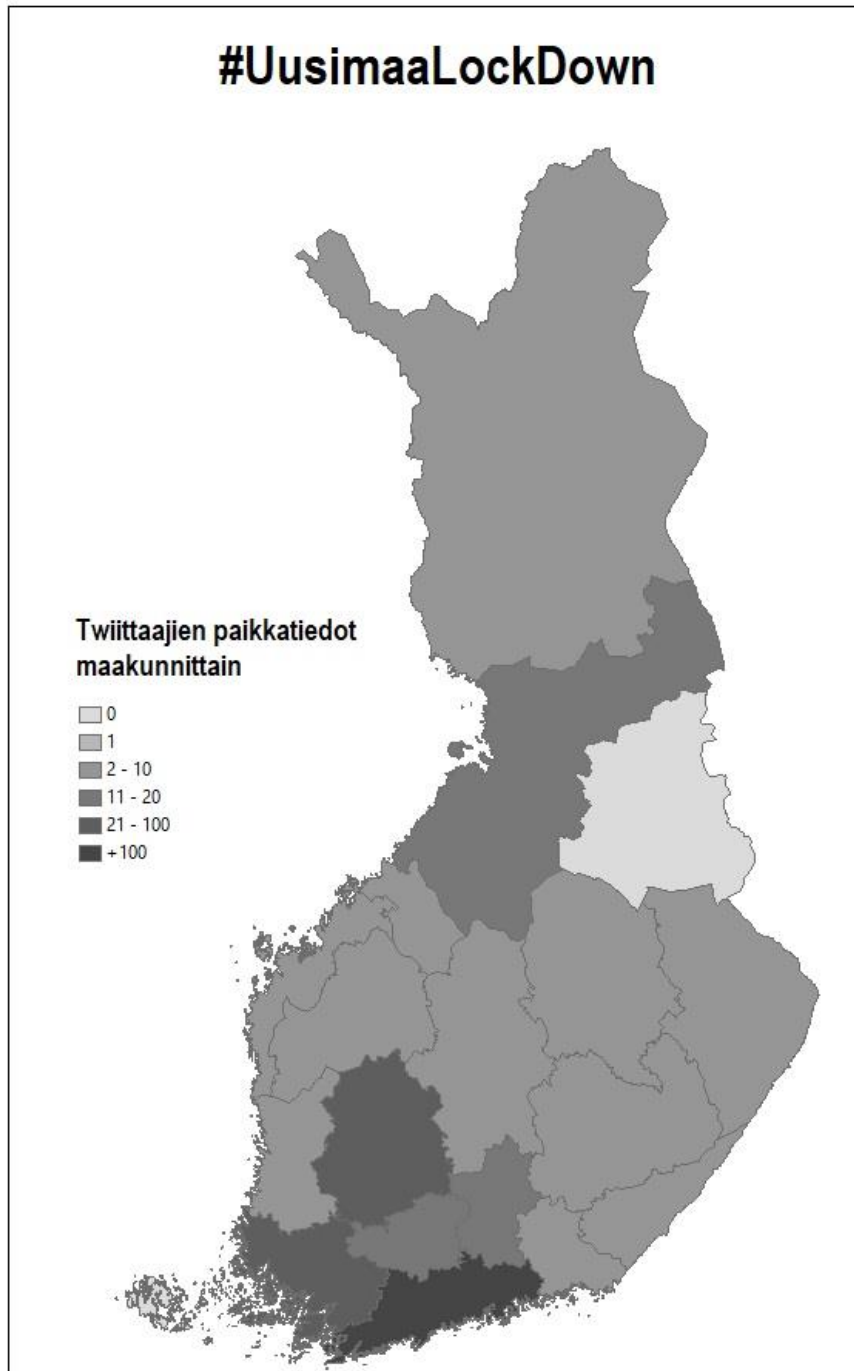
Wilson T. (2000). Human Information Behavior. *Informing Science* 3(2), 49-55. DOI: 10.28945/576

Ybarra, G. & Zimmerman, A. (2014). Online Aggression: The Influences of Anonymity and Social Modeling. *Psychology of Popular Media Culture*. DOI: 10.1037/ppm0000038.

Zinoviev, D. (2011). Information Diffusion in Social Media. *Social Networking and Community Behavior Modeling: Qualitative and Quantitative Measures* 146-163. DOI: 10.4018/978-1-61350-444-4

## 8 LIITTEET

Liite 1. Kartta, johon on visualisoitu #UusimaaLockDown twiittien määrät ja sijainnit. Kartta-aineistona on käytetty Maanmittauslaitoksen ”hallintorajat, teemakartoille, ei merialueita”-aineistoa, 2020 (CC 4.0 -lisenssi).



Liite 2. Piirakkadiagrammi #UusimaaLockDown -aihetunnistetta twiitanneiden käyttäjien paikkatiedot maakunnittain.

